



DUD. ...  
NAVAL ...  
MAY 1961











# NAVAL POSTGRADUATE SCHOOL

## Monterey, California



# THESIS

A STATISTICAL ANALYSIS OF MONTHLY RAINFALL FOR  
MONTEREY PENINSULA AND THE CARMEL VALLEY IN  
CENTRAL CALIFORNIA

by

David Frederick Davis  
March 1981

Thesis Advisor:

P.A. Jacobs

Approved for public release; distribution unlimited

T199165



| REPORT DOCUMENTATION PAGE  |                       | READ INSTRUCTIONS<br>BEFORE COMPLETING FORM                          |
|--|-----------------------|--|
| 1. REPORT NUMBER   | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER  |
| 4. TITLE (and Subtitle)<br>A Statistical Analysis of Monthly Rainfall<br>for Monterey Peninsula and the Carmel<br>Valley in Central California   |                       | 5. TYPE OF REPORT & PERIOD COVERED<br>Master's Thesis;<br>March 1981 |
| 7. AUTHOR(s)<br>David Frederick Davis  |                       | 6. PERFORMING ORG. REPORT NUMBER                                     |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Naval Postgraduate School<br>Monterey, California 93940   |                       | 8. CONTRACT OR GRANT NUMBER(s)                                       |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>Naval Postgraduate School<br>Monterey, California   |                       | 10. PROGRAM ELEMENT, PROJECT, TASK<br>AREA & WORK UNIT NUMBERS       |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)  |                       | 12. REPORT DATE<br>March 1981  |
|  |                       | 13. NUMBER OF PAGES<br>207   |
|  |                       | 15. SECURITY CLASS. (of this report)<br>Unclassified                 |
|  |                       | 15a. DECLASSIFICATION/DOWNGRADING<br>SCHEDULE                        |
| 16. DISTRIBUTION STATEMENT (of this Report)<br>Approved for public release; distribution unlimited   |                       |  |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)   |                       |  |
| 18. SUPPLEMENTARY NOTES  |                       |  |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number)<br>Monthly Rainfall Models                      Logistic Analysis<br>California Rainfall                              Logistic Time Series<br>Seasonal Rainfall Forecasting<br>ARMA Models  |                       |  |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number)<br>This thesis presents a statistical analysis of the monthly rainfall for the Monterey Peninsula and the Carmel Valley in Central California. The analysis begins with the simple first-order autoregressive Markov model, which is found to be weak. Next, 2x2 contingency tables are used to identify predictors, one of which is found to be January rainfall. Finally, logistic analysis is used to quantify the predictive ability |                       |  |



of January.

This paper attempts to analyze rainfall time series in the statistical sense. No attempt is made to provide a physical explanation of the findings from the point of view of a meteorologist.





Approved for public release; distribution unlimited

A Statistical Analysis of Monthly Rainfall for  
Monterey Peninsula and the Carmel Valley in  
Central California

by

David Frederick Davis  
Captain, United States Army  
B.S., Colorado School of Mines, 1972

Submitted in partial fulfillment of the  
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH



## ABSTRACT

This thesis presents a statistical analysis of the monthly rainfall for the Monterey Peninsula and the Carmel Valley in Central California. The analysis begins with the simple first-order autoregressive Markov model, which is found to be weak. Next, 2x2 contingency tables are used to identify predictors, one of which is found to be January rainfall. Finally, logistic analysis is used to quantify the predictive ability of January.

This paper attempts to analyze rainfall time series in the statistical sense. No attempt is made to provide a physical explanation of the findings from the point of view of a meteorologist.



# TABLE OF CONTENTS

|      |  |     |
|------|--|-----|
| I.   | INTRODUCTION -----                           | 20  |
|      | A. THE PROBLEM -----                         | 20  |
|      | B. NOTATION -----                            | 20  |
|      | C. METHODS OF ANALYSIS -----                 | 21  |
| II.  | THE DATA -----                               | 24  |
|      | A. GENERAL -----                             | 24  |
|      | B. DATA SET 'RN' -----                       | 26  |
|      | C. DATA SET 'FL' -----                       | 40  |
|      | D. DATA SET 'SC' -----                       | 52  |
| III. | FIRST ORDER MARKOV MODEL -----               | 65  |
|      | A. THEORY -----                              | 65  |
|      | B. DATA SET 'RN' -----                       | 67  |
|      | C. DATA SET 'FL' -----                       | 82  |
|      | D. DATA SET 'SC' -----                       | 97  |
| IV.  | VALIDATION OF FIRST ORDER MARKOV MODEL ----- | 114 |
|      | A. THEORY -----                              | 114 |
|      | B. DATA SET 'RN' -----                       | 115 |
|      | C. DATA SET 'FL' -----                       | 121 |
|      | D. DATA SET 'SC' -----                       | 126 |
|      | E. CONCLUSIONS -----                         | 130 |
| V.   | 2x2 TABLES -----                             | 131 |
|      | A. THEORY -----                              | 131 |
|      | B. ANALYSIS -----                            | 137 |
|      | C. OTHER RESULTS -----                       | 141 |



|                           |                                     |     |
|---------------------------|-------------------------------------|-----|
| VI.                       | LOGISTIC ANALYSIS -----             | 146 |
| A.                        | THEORY -----                        | 146 |
| B.                        | ANALYSIS -----                      | 150 |
| C.                        | DISCUSSION -----                    | 161 |
| VII.                      | VALIDATION OF LOGISTIC MODELS ----- | 169 |
| A.                        | GENERAL -----                       | 169 |
| B.                        | RESULTS -----                       | 172 |
| C.                        | DISCUSSION -----                    | 176 |
| VIII.                     | FURTHER FINDINGS -----              | 177 |
| A.                        | SUMMER MONTHS -----                 | 177 |
| B.                        | SIGNIFICANCE OF JANUARY -----       | 181 |
| IX.                       | SUMMARY -----                       | 185 |
| APPENDIX A:               | DATA SET RN -----                   | 186 |
| APPENDIX B:               | DATA SET FL -----                   | 191 |
| APPENDIX C:               | DATA SET SC -----                   | 198 |
| LIST OF REFERENCES        | -----                               | 205 |
| INITIAL DISTRIBUTION LIST | -----                               | 207 |





# LIST OF TABLES

|     |  |    |
|-----|--|----|
| 1.  | ESTIMATED AUTOCORRELATIONS OF YEARLY TOTAL RAINFALL<br>FOR DATA SET RN -----   | 29 |
| 2.  | MONTHLY MEANS AND VARIANCE FOR DATA SET RN -----   | 31 |
| 3.  | ESTIMATED AUTOCORRELATIONS OF MONTHLY RAINFALL<br>ANOMALIES FOR DATA SET RN -----  | 34 |
| 4.  | MONTHLY MEANS AND VARIANCE FOR LOGGED DATA SET RN----  | 36 |
| 5.  | ESTIMATED AUTOCORRELATIONS OF LOGGED ANOMALIES<br>OF MONTHLY RAINFALL FROM DATA SET RN -----   | 40 |
| 6.  | ESTIMATED AUTOCORRELATIONS OF YEARLY TOTAL RAINFALL<br>FOR DATA SET FL -----   | 43 |
| 7.  | MONTHLY MEANS AND VARIANCE FOR DATA SET FL -----   | 44 |
| 8.  | ESTIMATED AUTOCORRELATIONS OF MONTHLY RAINFALL<br>ANOMALIES FOR DATA SET FL -----  | 47 |
| 9.  | MONTHLY MEANS AND VARIANCE FOR LOGGED DATA SET FL----  | 48 |
| 10. | ESTIMATED AUTOCORRELATIONS OF LOGGED ANOMALIES FROM<br>MONTHLY RAINFALL OF DATA SET FL -----   | 51 |
| 11. | ESTIMATED AUTOCORRELATIONS OF YEARLY TOTAL RAINFALL<br>FOR DATA SET SC -----   | 55 |
| 12. | MONTHL MEANS AND VARIANCE FOR DATA SET SC -----  | 56 |
| 13. | ESTIMATED AUTOCORRELATIONS OF MONTHLY RAINFALL<br>ANOMALIES FOR DATA SET SC -----  | 59 |
| 14. | MONTHLY MEANS AND VARIANCE OF LOGGED DATA SET SC ----  | 60 |
| 15. | ESTIMATED AUTOCORRELATIONS OF LOGGED ANOMALIES OF<br>MONTHLY RAINFALL FROM DATA SET SC -----   | 64 |
| 16. | ESTIMATED PARTIAL-AUTOCORRELATIONS FOR LOGGED<br>RAINFALL ANOMALIES OF DATA SET RN -----   | 68 |
| 17. | ACTUAL AND EXPECTED NUMBER OF TURNING POINTS AND<br>ACTUAL AND EXPECTED PHASE FREQUENCIES FOR THE FIRST<br>ORDER MARKOV RESIDUALS FROM THE LOGGED RAINFALL<br>ANOMALIES OF DATA SET RN ----- | 72 |



|     |   |     |
|-----|---|-----|
| 18. | GENERAL STATISTICS OF FIRST ORDER MARKOV RESIDUALS<br>FROM LOGGED RAINFALL ANOMALIES OF DATA SET RN -----   | 73  |
| 19. | ACTUAL AND EXPECTED NUMBER OF TURNING POINTS AND<br>ACTUAL AND EXPECTED PHASE FREQUENCIES FOR THE<br>FIRST ORDER MARKOV RESIDUALS OF THE LOGGED RAINFALL<br>FOR WINTER MONTHLY ONLY, DATA SET RN -----            | 80  |
| 20. | GENERAL STATISTICS OF FIRST ORDER MARKOV RESIDUALS<br>FROM LOGGED RAINFALL ANOMALIES OF WINTER MONTHS,<br>DATA SET RN -----   | 81  |
| 21. | ESTIMATED PARTIAL AUTOCORRELATIONS FOR LOGGED<br>RAINFALL ANOMALIES OF DATA SET FL -----  | 83  |
| 22. | ACTUAL AND EXPECTED NUMBER OF TURNING POINTS AND<br>ACTUAL AND EXPECTED PHASE FREQUENCIES FROM THE<br>FIRST ORDER MARKOV RESIDUALS FROM DATA SET FL -----   | 87  |
| 23. | GENERAL STATISTICS OF FIRST ORDER MARKOV RESIDUALS<br>FROM LOGGED RAINFALL ANOMALIES OF DATA SET FL-----  | 87  |
| 24. | ACTUAL AND EXPECTED NUMBER OF TURNING POINTS AND<br>ACTUAL AND EXPECTED PHASE FREQUENCIES FROM THE<br>FIRST ORDER MARKOV RESIDUALS OF WINTER MONTHLY<br>ONLY, DATA SET FL -----                                   | 95  |
| 25. | GENERAL STATISTICS OF FIRST ORDER MARKOV RESIDUALS<br>FROM LOGGED RAINFALL ANOMALIES OF WINTER MONTHS<br>ONLY, DATA SET FL -----  | 95  |
| 26. | ESTIMATED PARTIAL AUTOCORRELATIONS FOR LOGGED<br>RAINFALL ANOMALIES OF DATA SET SC -----  | 98  |
| 27. | ACTUAL AND EXPECTED NUMBER OF TURNING POINTS AND<br>ACTUAL AND EXPECTED PHASE FREQUENCIES FROM THE<br>FIRST ORDER MARKOV RESIDUALS OF DATA SET SC -----   | 102 |
| 28. | GENERAL STATISTICS OF FIRST ORDER MARKOV RESIDUALS<br>FROM LOGGED RAINFALL ANOMALIES OF DATA SET SC -----   | 102 |
| 29. | ACTUAL AND EXPECTED NUMBER OF TURNING POINTS AND<br>ACTUAL AND EXPECTED PHASE FREQUENCIES FROM THE<br>FIRST ORDER MARKOV RESIDUALS OF THE LOGGED RAINFALL<br>ANOMALIES OF THE WINTER MONTHS ONLY, DATA SET SC---- | 111 |
| 30. | GENERAL STATISTICS OF FIRST ORDER MARKOV RESIDUALS<br>FROM LOGGED RAINFALL ANOMALIES OF WINTER MONTHS<br>ONLY, DATA SET SC -----  | 112 |



|     |  |     |
|-----|--|-----|
| 31. | ACTUAL AND EXPECTED NUMBER OF TURNING POINTS AND<br>ACTUAL AND EXPECTED PHASE FREQUENCIES FOR THE<br>FORECAST ERRORS OF THE FIRST ORDER MARKOV MODEL<br>APPLIED TO THE WINTER MONTHS OF RESERVED DATA<br>SET RN -----      | 118 |
| 32. | GENERAL STATISTICS OF FORECAST ERRORS FROM THE FIRST<br>ORDER MARKOV MODEL APPLIED TO THE WINTER MONTHS OF<br>RESERVED DATA SET RN -----   | 119 |
| 33. | ACTUAL AND EXPECTED NUMBER OF TURNING POINTS AND<br>ACTUAL AND EXPECTED PHASE FREQUENCIES FOR THE<br>FORECAST ERRORS OF THE FIRST ORDER MARKOV MODEL<br>APPLIED TO THE WINTER MONTHS OF RESERVED DATA<br>SET FL -----      | 123 |
| 34. | GENERAL STATISTICS OF FORECAST ERRORS FROM THE<br>FIRST ORDER MARKOV MODEL APPLIED TO THE WINTER<br>MONTHS OF RESERVED DATA SET FL -----   | 124 |
| 35. | ACTUAL AND EXPECTED NUMBER OF TURNING POINTS AND<br>ACTUAL AND EXPECTED PHASE FREQUENCIES FOR THE<br>FORECAST ERRORS OF THE FIRST ORDER MARKOV MODEL<br>APPLIED TO THE WINTER MONTHS ONLY OF RESERVED<br>DATA SET SC ----- | 128 |
| 36. | GENERAL STATISTICS OF FORECAST ERRORS FROM THE<br>FIRST ORDER MARKOV MODEL APPLIED TO THE WINTER<br>MONTHS OF RESERVED DATA SET SC -----   | 128 |
| 37. | SIGNIFICANCE OF OBSERVED DEPARTURES FROM THE<br>INDEPENDENCE OF SINGLE MONTH CONTROLS VERSUS<br>SUCCEEDING ELEVEN MONTH COMPLEMENTS -----  | 138 |
| 38. | SIGNIFICANCE OF OBSERVED DEPARTURES FROM INDEPENDENCE<br>OF PAIRED MONTHS CONTROLS VERSUS SUCCEEDING TEN MONTH<br>COMPLEMENTS -----  | 139 |
| 39. | SIGNIFICANCE OF OBSERVED DEPARTURES FROM INDEPENDENCE<br>OF TRIPLES OF MONTHS CONTROLS VERSUS SUCCEEDING NINE<br>MONTH COMPLEMENTS -----   | 139 |
| 40. | SIGNIFICANCE OF OBSERVED DEPARTURES FROM INDEPENDENCE<br>OF FOUR-TUPLES OF MONTHS CONTROLS VERSUS SUCCEEDING<br>EIGHT MONTH COMPLEMENTS -----  | 140 |
| 41. | ODDS RATIO OF JANUARY VERSUS FEBRUARY THROUGH DECEMBER<br>AND JANUARY PLUS DECEMBER VERSUS FEBRUARY THROUGH<br>NOVEMBER -----  | 141 |
| 42. | ANOVA FOR REGRESSION OF SIMPLE LINEAR MODEL FOR<br>ALL DATA SETS -----   | 142 |



|     |   |     |
|-----|---|-----|
| 43. | ANOVA FOR REGRESSION OF SIMPLE LINEAR MODEL WITH<br>MEANS REMOVED FOR ALL DATA SETS -----                                       | 144 |
| 44. | DATA SET RN, LOGGED JANUARY ANOMALIES AND SUCCESSES<br>FOR GROUPED AND UNGROUPED FORMS -----                                    | 151 |
| 45. | DATA SET FL, LOGGED JANUARY ANOMALIES AND SUCCESSES<br>FOR GROUPED AND UNGROUPED FORMS -----                                    | 152 |
| 46. | DATA SET SC, LOGGED JANUARY ANOMALIES AND SUCCESSES<br>FOR GROUPED AND UNGROUPED FORMS -----                                    | 153 |
| 47. | ORDINARY LEAST SQUARES REGRESSION WITH THE MODEL<br>OF EQUATION VI.7 -----  | 154 |
| 48. | ITERATIVELY REWEIGHTED LEAST SQUARES REGRESSION<br>USING BIWEIGHTS FOR THE MODEL OF EQUATION VI.7 -----                         | 158 |
| 49. | MAXIMUM LIKELIHOOD ESTIMATES OF $\alpha$ AND $\beta$ ALONG<br>WITH ESTIMATES OF THEIR VARIANCE FOR ALL THREE<br>DATA SETS ----- | 159 |
| 50. | PARAMETER FIT RECAPITULATION FOR ALL DATA SETS -----  | 163 |
| 51. | DATA USED FOR MODEL FITS OF LOGISTIC MODELS -----   | 167 |
| 52. | RESERVED DATA IN FORM FOR THE LOGISTIC ANALYSIS-----  | 170 |
| 53. | RESULTS OF LOGISTIC VALIDATION ON DATA SET RN -----   | 172 |
| 54. | RESULTS OF LOGISTIC VALIDATION ON DATA SET FL -----   | 174 |
| 55. | RESULTS OF VALIDATION ON DATA SET SC -----  | 175 |







# LIST OF FIGURES

|     |   |    |
|-----|---|----|
| 1.  | Location of rainfall data sets and the years available -----                | 25 |
| 2.  | Monthly rainfall in inches for data set RN -----                            | 27 |
| 3.  | Yearly total rainfall for data set RN -----<br>(1951-1974)                  | 28 |
| 4.  | Correlogram of yearly total rainfall for data set RN -----                  | 29 |
| 5.  | Lag one plot of yearly rainfall data for data set RN -----                  | 30 |
| 6.  | Monthly means for data set RN -----   | 32 |
| 7.  | Monthly rainfall anomalies in inches for data set RN -----                  | 33 |
| 8.  | Correlogram of the month rainfall anomalies for data set RN -----           | 34 |
| 9.  | Plot of monthly variance against monthly means for data set RN -----        | 35 |
| 10. | Monthly means of logged data set RN -----                                   | 37 |
| 11. | Plot of monthly variance against monthly means for logged data set RN ----- | 37 |
| 12. | Logged anomalies of monthly rainfall for data set RN -----                  | 38 |
| 13. | Correlogram of logged anomalies of monthly rainfall from data set RN -----  | 39 |
| 14. | Rainfall in inches, by month, of data set FL -----                          | 41 |
| 15. | Yearly total rainfall for data set FL -----<br>(1937 - 1974)                | 42 |
| 16. | Correlogram of yearly total rainfall for data set FL -----                  | 43 |
| 17. | Monthly means for data set FL -----   | 44 |
| 18. | Monthly rainfall anomalies in inches for data set FL -----                  | 45 |
| 19. | Correlogram of monthly rainfall anomalies for data set FL -----             | 46 |



|     |   |    |
|-----|---|----|
| 20. | Plot of monthly variance against monthly means<br>for data set FL -----   | 47 |
| 21. | Monthly means of logged data set FL -----   | 48 |
| 22. | Plot of monthly variance against monthly means<br>for logged data set FL -----  | 49 |
| 23. | Months of logged rainfall anomalies for data<br>set FL -----  | 49 |
| 24. | Correlogram of logged anomalies of monthly<br>rainfall from data set FL -----   | 51 |
| 25. | Monthly plot of rainfall in inches<br>for data set SC -----   | 52 |
| 26. | Yearly total rainfall for data set SC -----   | 54 |
| 27. | Correlogram of yearly total rainfall for<br>data set SC -----   | 55 |
| 28. | Monthly means for data set SC -----   | 56 |
| 29. | Rainfall anomalies in inches for data set SC -----  | 57 |
| 30. | Correlogram of monthly rainfall anomalies for<br>data set SC -----  | 59 |
| 31. | Plot of monthly variance against monthly means<br>for data set SC -----   | 60 |
| 32. | Monthly means of logged data set SC -----   | 61 |
| 33. | Plot of monthly variance against monthly means<br>for logged data set SC -----  | 61 |
| 34. | Months of logged rainfall anomalies from data<br>set SC -----   | 62 |
| 35. | Correlogram of logged anomalies of monthly<br>rainfall from data set SC -----   | 64 |
| 36. | Partial correlogram of the logged rainfall<br>anomalies of data set RN -----  | 67 |
| 37. | First order Markov residuals from logged rainfall<br>anomalies of data set RN -----   | 69 |
| 38. | Autocorrelations of residuals from first order Markov<br>process applied to the logged rainfall anomalies of<br>data set RN ----- | 71 |



|     |  |    |
|-----|--|----|
| 39. | Lag one plot of first order Markov residuals from<br>logged rainfall anomalies of data set RN -----  | 71 |
| 40. | First order Markov residuals versus lag one<br>data point from logged rainfall anomalies of<br>data set RN -----                           | 72 |
| 41. | Standardized normal plot of first order Markov<br>residuals from logged rainfall anomalies data<br>set RN -----                            | 73 |
| 42. | Histogram of first order Markov residuals from<br>logged rainfall anomalies of data set RN -----   | 74 |
| 43. | Winter months only of logged rainfall anomalies<br>of data set RN -----  | 75 |
| 44. | Correlogram of winter months only of logged<br>rainfall anomalies from data set RN -----   | 76 |
| 45. | Partial autocorrelations of winter months only of<br>logged rainfall anomalies from data set RN -----                                      | 76 |
| 46. | First order Markov residuals of logged rainfall<br>anomalies for winter months only of data set RN ----                                    | 77 |
| 47. | Correlogram of first order Markov residuals of<br>logged rainfall anomalies for winter months only<br>of data set RN -----                 | 79 |
| 48. | Lag one plot of first order Markov residuals from<br>logged rainfall anomalies for winter months only,<br>data set RN -----                | 79 |
| 49. | First order Markov residuals versus lag one data<br>point from logged rainfall anomalies of winter<br>months only, data set RN -----       | 80 |
| 50. | Standardized normal plot of first order Markov<br>residuals from logged rainfall anomalies of<br>winter months only from data set RN ----- | 81 |
| 51. | Histogram of first order Markov residuals from<br>logged rainfall anomalies of winter months only,<br>data set RN -----                    | 82 |
| 52. | Partial correlogram of the logged rainfall<br>anomalies for data set FL -----  | 83 |
| 53. | First order Markov residual from logged rainfall<br>anomalies of data set FL -----   | 84 |



|     |  |    |
|-----|--|----|
| 54. | Autocorrelations of residuals from first order Markov process applied to the logged rainfall anomalies of data set FL -----      | 85 |
| 55. | Lag one plot of first order Markov residuals from logged rainfall anomalies of data set FL -----                                 | 86 |
| 56. | First order Markov residuals versus lag one data points from logged rainfall anomalies of data set FL -----                      | 86 |
| 57. | Standardized normal plot of first order Markov residuals from logged rainfall anomalies of data set FL -----                     | 88 |
| 58. | Histogram of first order Markov residuals from logged rainfall anomalies of data set FL -----                                    | 88 |
| 59. | Winter months only of logged rainfall anomalies of data set FL -----   | 89 |
| 60. | Correlogram of winter months only, logged rainfall anomalies from data set FL -----  | 91 |
| 61. | Partial correlogram of winter months only, logged rainfall anomalies from data set FL -----                                      | 91 |
| 62. | First order Markov residuals of logged rainfall anomalies for winter months only, data set FL -----                              | 92 |
| 63. | Correlogram of first order Markov residuals of logged rainfall anomalies from winter months only, data set FL -----              | 93 |
| 64. | Lag one plot of first order Markov residuals from logged rainfall anomalies of winter months only, data set FL -----             | 94 |
| 65. | First order Markov residuals versus lag one data point from logged rainfall anomalies of winter months only, data set FL -----   | 94 |
| 66. | Standardized normal plot of first order Markov residuals from logged rainfall anomalies of winter months only, data set FL ----- | 96 |
| 67. | Histogram of first order Markov residuals from logged rainfall anomalies of winter months only, data set FL -----                | 96 |
| 68. | Partial correlogram of the logged anomalies for data set SC -----  | 97 |





|     |  |     |
|-----|--|-----|
| 69. | First order Markov residuals from logged rainfall anomalies of data set SC -----   | 98  |
| 70. | Autocorrelations of residuals from first order Markov process applied to the logged rainfall anomalies of data set SC -----      | 100 |
| 71. | Lag one plot of first order Markov residuals from logged rainfall anomalies of data set SC -----                                 | 101 |
| 72. | First order Markov residuals versus lag one data point from logged rainfall anomalies of data set SC -----                       | 101 |
| 73. | Standardized normal plot of first order Markov residuals from logged rainfall anomalies of data set SC -----                     | 103 |
| 74. | Histogram of first order Markov residuals from logged rainfall anomalies of data set SC -----                                    | 103 |
| 75. | Winter months only, logged rainfall anomalies of data set SC -----   | 104 |
| 76. | Correlogram of winter months only, logged rainfall anomalies from data set SC -----  | 106 |
| 77. | Partial correlogram of winter months only, logged rainfall anomalies from data set SC -----                                      | 107 |
| 78. | First order Markov residuals of logged rainfall anomalies, for winter months only, data set SC -----                             | 108 |
| 79. | Correlogram of first order Markov residuals of logged rainfall anomalies from winter months only, data set SC -----              | 110 |
| 80. | Lag one plot of first order Markov residuals from logged rainfall anomalies of winter months only, data set SC -----             | 110 |
| 81. | First order Markov residuals versus lag one data point form logged rainfall anomalies of winter months only, data set SC -----   | 111 |
| 82. | Standardized normal plot of first order Markov residuals from logged rainfall anomalies of winter months only, data set SC ----- | 112 |
| 83. | Histogram of first order Markov residuals from logged rainfall anomalies of winter months only, data set SC -----                | 113 |



|     |  |     |
|-----|--|-----|
| 84. | Reserved rainfall data for data set RN -----   | 115 |
| 85. | Logged rainfall anomalies of reserved<br>data set RN -----   | 116 |
| 86. | Forecast errors from first order Markov model<br>applied to winter months of reserved data set RN----  | 117 |
| 87. | Correlogram of forecast errors from first order<br>Markov model applied to winter months of reserved<br>data set RN -----                      | 117 |
| 88. | Standardized normal plot of forecast errors from<br>the first order Markov model applied to the winter<br>months of reserved data set RN ----- | 119 |
| 89. | Histogram of forecast errors from the first order<br>Markov model applied to the winter months of<br>reserved data set RN -----                | 120 |
| 90. | Reserved rainfall data for data set FL -----   | 121 |
| 91. | Logged rainfall anomalies of reserved data set FL---   | 122 |
| 92. | Forecast errors from the first order Markov model<br>applied to the winter months of reserved data<br>set FL -----                             | 122 |
| 93. | Correlogram of forecast errors from first order<br>Markov model applied to the winter months of<br>reserved data set FL -----                  | 123 |
| 94. | Standardized normal plot of forecast errors from<br>the first order Markov model applied to the winter<br>months of reserved data set FL ----- | 124 |
| 95. | Histogram of forecast errors from the first order<br>Markov model applied to the winter months of<br>reserved data set FL -----                | 125 |
| 96. | Reserved rainfall for data set SC -----  | 126 |
| 97. | Logged anomalies of reserved data set SC -----   | 126 |
| 98. | Forecast errors from first order Markov model<br>applied to the winter months of reserved data<br>set SC -----                                 | 127 |
| 99. | Correlogram of forecast errors from first order<br>Markov model applied to the winter months of<br>reserved data set SC -----                  | 127 |



|      |   |     |
|------|---|-----|
| 100. | Standardized normal plot of forecast errors from the first order Markov model applied to the winter months of reserved data set SC -----                | 129 |
| 101. | Histogram of forecast errors from the first order Markov model applied to the winter months of reserved data set SC -----                               | 129 |
| 102. | Typical 2x2 contingency Table -----   | 132 |
| 103. | Contours of log likelihood function for data set RN -----   | 160 |
| 104. | Contours of log likelihood function for data set FL -----   | 160 |
| 105. | Contours of log likelihood function for data set SC -----   | 161 |
| 106. | Estimated probability of greater-than-average total rest-of-year rainfall versus the anomaly of logged rainfall for January for data set RN ----        | 164 |
| 107. | Estimated probability of greater-than-average total rest-of-year rainfall versus the anomaly of logged rainfall for January for data set FL----         | 165 |
| 108. | Estimated probability of greater-than-average total rest-of-year rainfall versus the anomaly of logged rainfall for January for data set SC----         | 166 |
| 109. | 2x2 contingency Tables of reserved data controlled by the anomaly of January rainfall. The complement is the anomaly of the rest of year rainfall ----- | 171 |
| 110. | Plot of versus complement anomalies for data set RN -----   | 173 |
| 111. | Plot of versus complement anomalies for data set FL -----   | 174 |
| 112. | Plot of versus complement anomalies for maximum likelihood and IRWLS parameters for data set SC----   | 176 |
| 113. | Monthly plot of summer months only, means removed, data set RN -----  | 178 |
| 114. | Monthly plot of summer months only, means removed, data set FL -----  | 178 |
| 115. | Monthly plot of summer months only, means removed, data set SC -----  | 179 |



|      |  |     |
|------|--|-----|
| 116. | Yearly plot of total summer rainfall for<br>data set RN -----                          | 179 |
| 117. | Yearly plot of total summer rainfall for<br>data set FL -----                          | 180 |
| 118. | Yearly plot of total summer rainfall for<br>data set SC -----                          | 180 |
| 119. | Log-odds and significance versus additional<br>months cumulated through the year ----- | 183 |
| 120. | Log-odds and significance versus additional<br>months of the fall -----                | 184 |





## ACKNOWLEDGEMENT

I would like to thank my advisors, Professors D. P. Gaver and P. A. Jacobs for their guidance. My second reader, Professor R. J. Renard has made sure that no laws of meteorology have been destroyed, for which I am grateful.

Other who assisted me should also be mentioned, especially Professor A. L. Schoenstadt who allowed me access to the HP 9845B computer which drew all the figures. The personnel at the Monterey Peninsular water Management District, Dr. J. Williams, Mr. B. Buel, and Mr. K. Walsh, were also very helpful.



## I. INTRODUCTION

### A. THE PROBLEM

The Monterey Peninsula Water Management District, in Central California coastal area has as one of its responsibilities the duty to recommend and/or impose water rationing on its constituents. To do this in a rational way requires the District to have some formula for predicting future water availability. Although the techniques of modern meteorology are becoming more sophisticated and exact there is still the inability to make good long-range predictions. This thesis analyzes three series of Monterey County monthly rainfall data by purely statistical methodology in order to identify possible predictive formulas.

### B. NOTATION

Rainfall will be denoted as  $R_{t,m}$  which will represent inches of rain recorded for the  $t^{th}$  year and the  $m^{th}$  month. The year to be used is the California Water Year which begins in October and ends the following September. Thus  $R_{1,1}$  is the monthly rainfall for October of year, '1' and  $R_{6,8}$  is the monthly rainfall of May of year '6'.

An overstruck bar as in  $\bar{R}..$  will indicate the arithmetic average of a variable; in this case it is the arithmetic average of all years and months of rainfall.  $\bar{R}_m$  is the average of rainfall over the years for month  $m$ ;  $\bar{R}_t$ .



represents the yearly average for year  $t$ .

### C. METHODS OF ANALYSIS

Three methods were used to analyze the data. The first method was to model the series using autoregressive moving averages as described in Box and Jenkins [Ref. 1]. The second was to use 2x2 contingency tables to identify possible predictors. The third was logistic regression to quantify the findings of the 2x2 contingency table analysis. These three methods will be described in further sections of this paper.

#### 1. ARMA(p,q) Models

A widely used approach to time series modeling proposed by Box and Jenkins is the ARMA(p,q) model. This model is actually a joining of two types of model, the autoregressive and the moving average.

In the notation of Box and Jenkins:

let  $\{Z_t, t=1,2,\dots,n\}$  be a time series, then an ARMA(p,q) process may be written as:

I.1

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \dots + \phi_p \tilde{Z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

the  $\{a_t, t=1,2,\dots,n\}$  are assumed to be random shocks distributed as independent and identically distributed (iid)

random variables with mean zero and variance  $\sigma_a^2$  and

$\tilde{Z}_t = Z_t - \bar{Z}$ . The further assumption of normality is also usually made.

For purposes of this paper, a mapping of  $R_{t,m}$  into  $Z_r, r=l2(t-1)+m$  was made, and an ARMA analysis was



then conducted on this index transformed series. This analysis is described in section III.

## 2. 2x2 Table Analysis

In the validation of section IV it is found that the ARMA model is not very successful in describing the data. In section V the data is analyzed by means of 2x2 contingency tables. These tables are good tools for exploratory data analysis in that they provide a visual display of the data. Statistical procedures based on the null hypothesis of independence can be used to quantify the departure from independence. The theory of 2x2 tables, and contingency tables in general may be found in Fleiss [Ref. 3], Dixon and Massey [Ref. 5], Brownlee [Ref. 6], and Mood, Graybill, and Boes [Ref. 7].

For this paper, the contingency table approach is used to identify a month or group of months of a year whose rainfall can serve as a predictor for the rainfall during the remaining months of the year. One predictor that was suggested is the rainfall in the month of January.

## 3. Logistic Analysis

Once a predictor is tentatively identified it becomes necessary to quantify the degree, direction and accuracy of the predictor.

A logistic analysis is conducted by dividing the data for a year into two sets, the predictor or control set, and predictand or complement set. For this analysis, the predictor





is the logged anomaly of January rainfall for the year;  
that is, if  $X_t$  denotes the predictor or control for year  
 $t$ , then

$$X_t = \ln(R_{t,4}) - \frac{1}{N} \sum_{t=1}^N \ln(R_{t,4}) \quad \text{I.2}$$

(The logarithm is used to better symmetrize the model.) The  
complement is the raw anomaly of the total rainfall for the  
immediately subsequent eleven months; that is, if  $Y'_t$  denotes  
the complement for year  $t$ , then;

$$Y'_t = \left( \sum_{m=5}^{12} R_{t,m} + \sum_{m=1}^3 R_{t+1,m} \right) - \frac{1}{N-1} \sum_{t=1}^{N-1} \left( \sum_{m=5}^{12} R_{t,m} + \sum_{m=1}^3 R_{t+1,m} \right) \quad \text{I.3}$$

Finally, the data are transformed into a binary representa-  
tion, relative to zero as;

$$Y_t = \begin{cases} 0 & \text{if } Y'_t < 0 \\ 1 & \text{if } Y'_t > 0 \end{cases} \quad \text{I.4}$$

In section VI the model fit is

$$P(Y=1|X=x) = \frac{e^{\alpha + \beta x}}{1 + e^{\alpha + \beta x}} \quad \text{I.5}$$

Where  $x$  is as before and  $P(Y=1|X=x)$  is interpreted as:  
"the conditional probability that the subsequent eleven month  
total rainfall will be above its mean, given that the logged  
anomaly of January rainfall was ' $x$ '".



## II. THE DATA

### A. GENERAL

Three data sets were used for this analysis. The location at which these data sets were gathered is shown in Figure 1. As the figure indicates, two of the data sets are on the Monterey Peninsula proper, while the third set, SC, represents the Carmel River Watershed at the San Clemente Dam.

Although data exists in all cases to the present, all three sets were truncated at September of 1974. The remaining data, up through September of 1980 was reserved for validation of the models and methodology.

The data coordinates are:

Data set RN:     $36^{\circ} 35' 42''$  North Latitude  
                  $121^{\circ} 54' 43''$  West Longitude

Data set FL:     $36^{\circ} 35' 30''$  North Latitude  
                  $121^{\circ} 56' 30''$  West Longitude

Data set SC:     $36^{\circ} 26' 12''$  North Latitude  
                  $121^{\circ} 42' 30''$  West Longitude



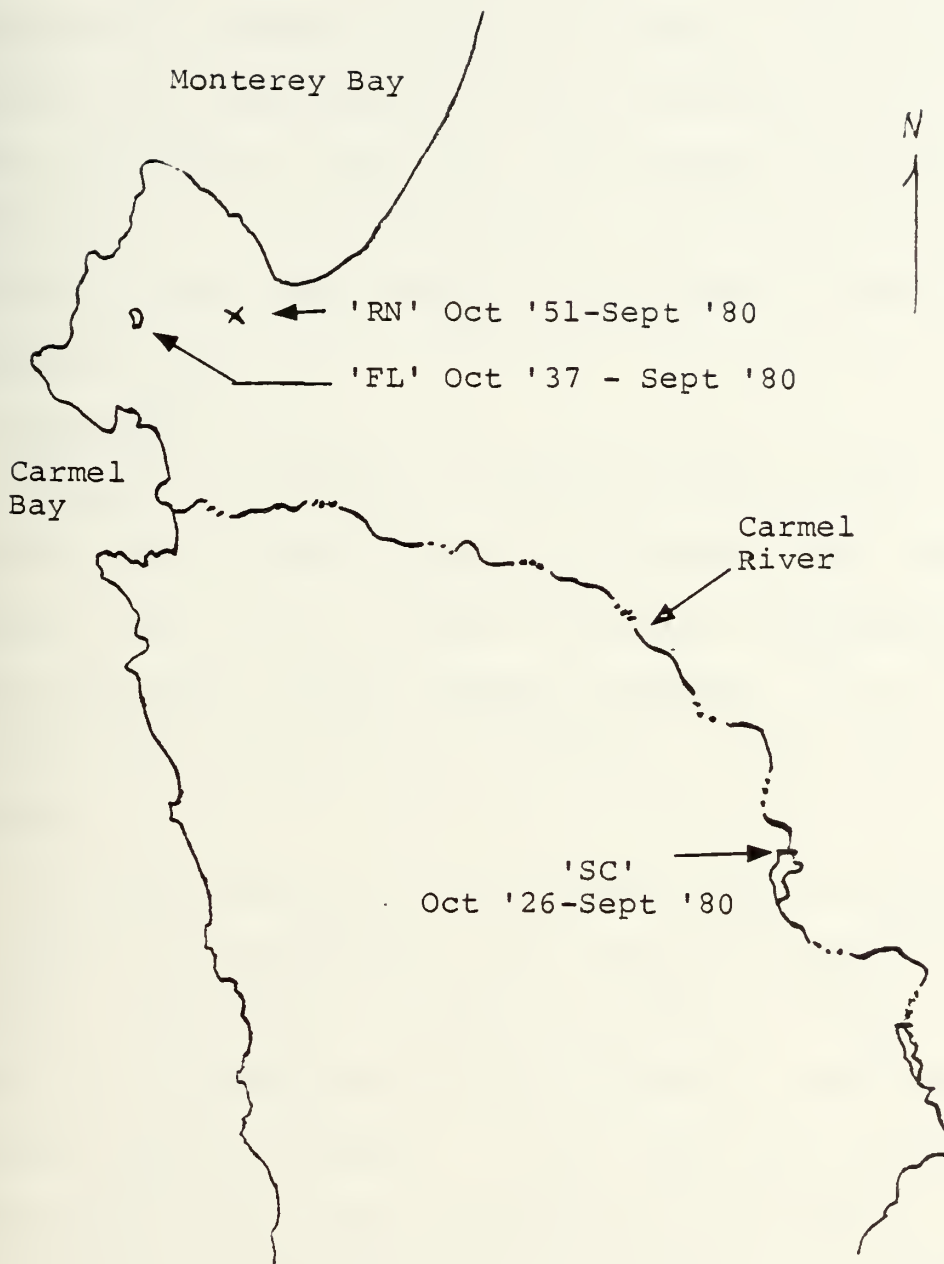


Figure 1. Location of rainfall data sets and the years available.



## B. DATA SET RN

Data set RN consists of monthly rainfall amounts gathered by Professor R.J. Renard, Cooperative Observer for the National Weather Service Climatological Station, Monterey, California. The data set begins in June 1951 and currently terminates in September 1980. As was stated above, the analysis was conducted only on that data between and including October 1951 and September 1974.

### 1. Raw Data

Appendix A contains a listing of data set RN. Figure 2 shows the raw data set. Month 1 is October 1951, month 148 is January 1964, and up to month 288 which is September 1974. As can be seen the data are strongly seasonal. This is enough to indicate that the series, as stated, is highly non-stationary.

The data presented so far deals with only monthly data. Next to be considered is the series of yearly total rainfalls. The results are shown in Figure 3 (Yearly total rainfall), 4 (Correlogram of yearly rainfall), and Table 1 (Estimated Autocorrelations). In this case, the correlogram indicates stationarity and independence of the yearly series. A plot of the lag one relationships, Figure 5, reinforces this indication of independence.





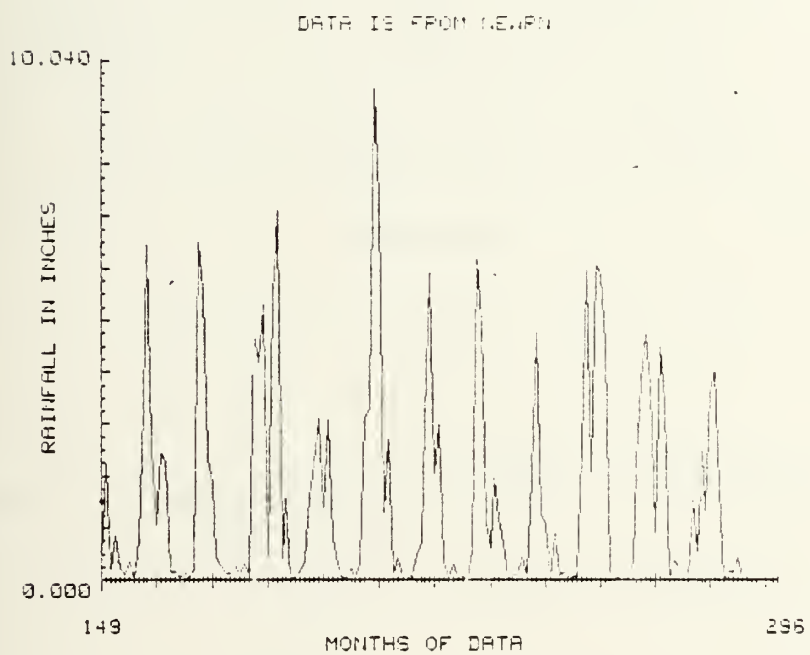
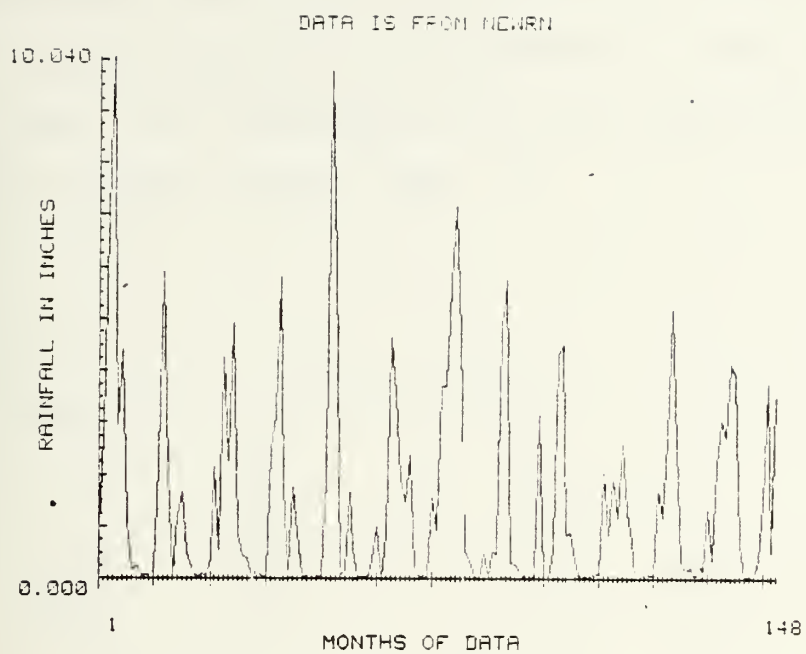


Figure 2. Monthly rainfall in inches for data set RN.



The correlograms and Partial Correlograms to follow indicate the 95% approximate significance levels using dashed lines. For development of these significance levels see Box and Jenkins [Ref. 1].

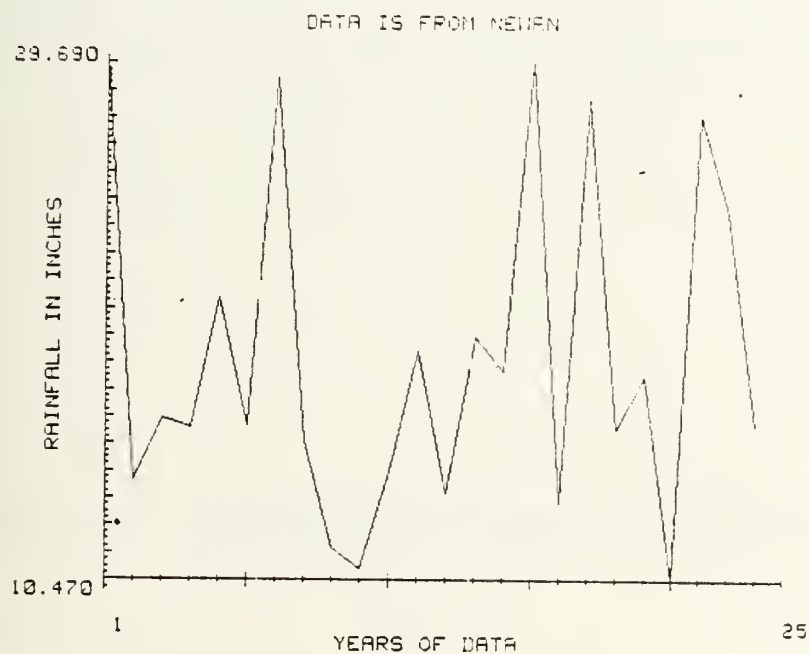


Figure 3. Yearly total rainfall for data set RN (1951 - 1974).



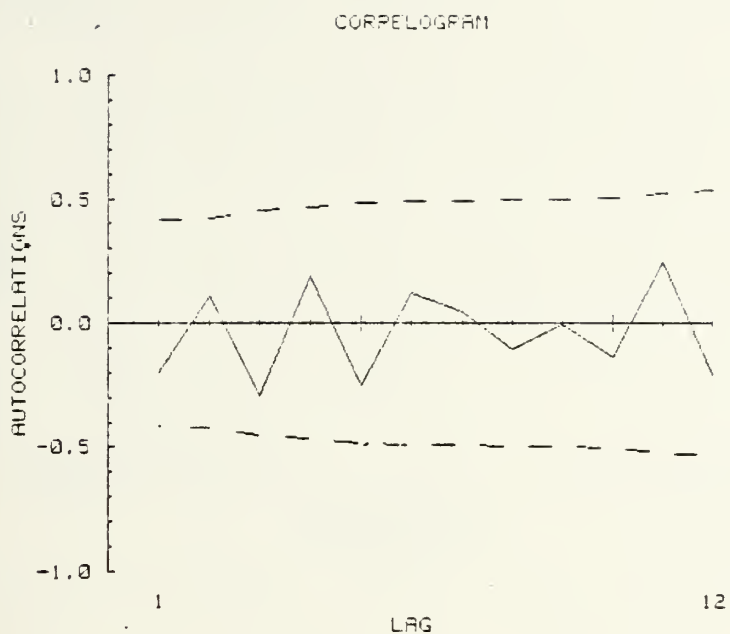


Figure 4. Correlogram of Yearly total rainfall for data set RN.

TABLE 1

ESTIMATED AUTOCORRELATIONS OF YEARLY  
TOTAL RAINFALL FOR DATA SET RN

AUTOCORRELATIONS

| LAG | VALUE | LAG | VALUE |
|-----|-------|-----|-------|
| 1   | -.200 | 7   | .044  |
| 2   | .109  | 8   | -.109 |
| 3   | -.295 | 9   | -.009 |
| 4   | .186  | 10  | -.139 |
| 5   | -.249 | 11  | .248  |
|     |       | 12  | -.211 |



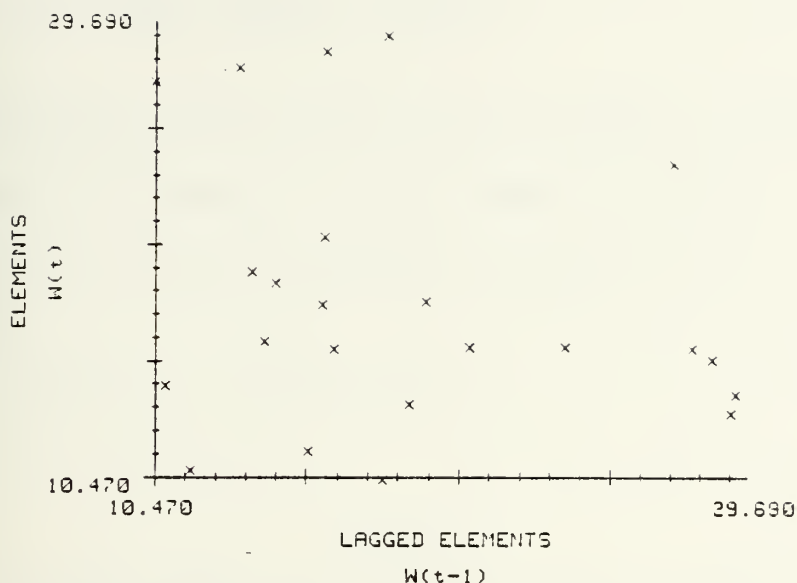


Figure 5. Lag one plot of yearly rainfall data for data set RN.

## 2. Swept Data

Pierce [Ref. 9] and Hipel [Ref. 11] suggest various ways to remove the seasonality of data sets like RN, FL, and SC. The basic, and most straight forward of these methods is to remove the various monthly means. This is accomplished by the following replacement:

let

$$\tilde{R}_{t,m} = R_{t,m} - \bar{R}_m \quad \text{II.5}$$

where  $\bar{R}_m$  represents the mean of the month  $m$ .

One statistic that is a byproduct of the calculations of  $\bar{R}_m$  is  $S^2_m$  defined as the estimated variance of the





monthly data points:

$$S_m^2 = \frac{1}{N-1} \left( \sum_{t=1}^N R_{t,m}^2 - N\bar{R}_m^2 \right); \quad \text{II.6}$$

These statistics for data set RN are shown in Table 2, and illustrated in Figure 6. In the same way as the raw data mapped into a series, a series is created from

$\tilde{R}_{t,m}$  as;

$$\tilde{Z}_r = \tilde{R}_{t,m}, r=12(t-1)+m \quad \text{II.7}$$

TABLE 2

MONTHLY MEANS AND VARIANCE FOR DATA SET RN

| MONTH | MEAN  | VARIANCE |
|-------|-------|----------|
| 1     | .677  | .5292    |
| 2     | 2.478 | 3.6354   |
| 3     | 3.355 | 5.9147   |
| 4     | 4.124 | 5.4146   |
| 5     | 2.592 | 4.5122   |
| 6     | 2.639 | 3.7414   |
| 7     | 1.708 | 2.7699   |
| 8     | .434  | .2494    |
| 9     | .219  | .1004    |
| 10    | .058  | .0056    |
| 11    | .104  | .0129    |
| 12    | .275  | .3953    |



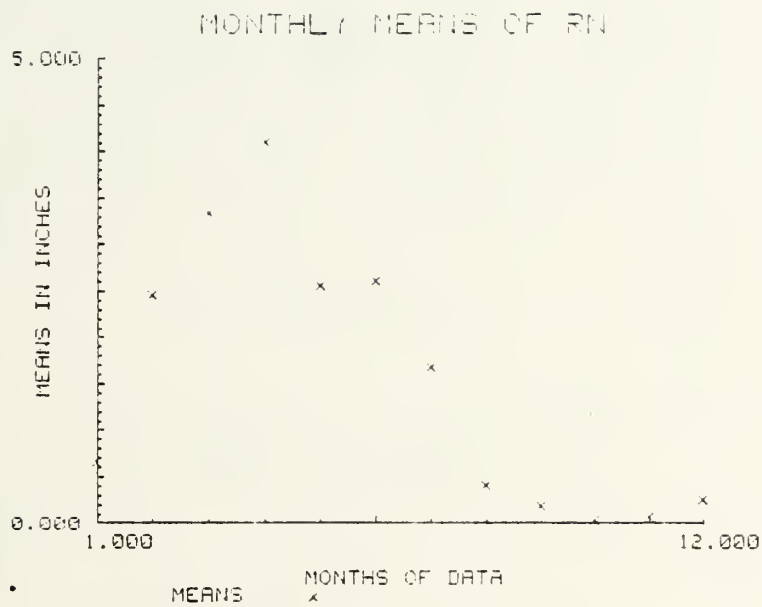


Figure 6. Monthly means for data set RN



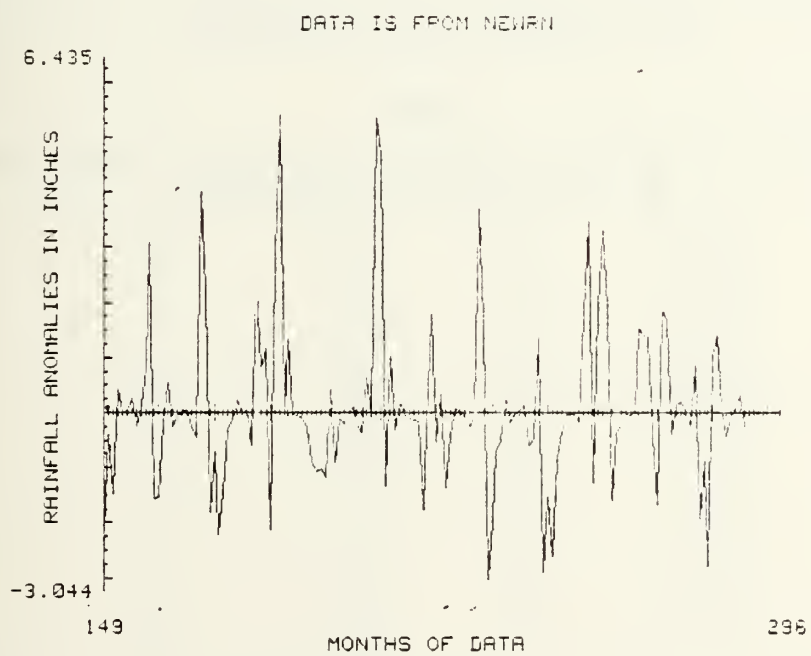
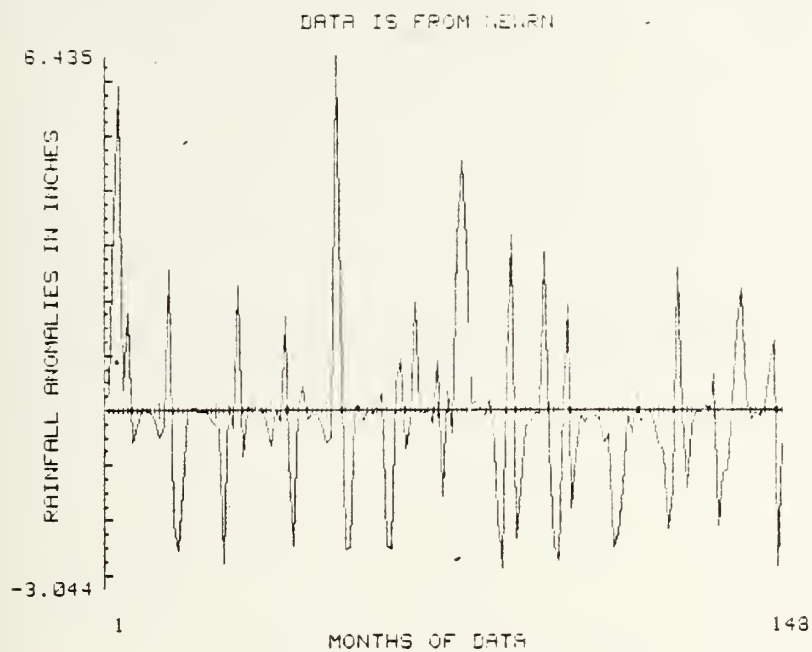


Figure 7. Monthly rainfall anomalies in inches for data set RN



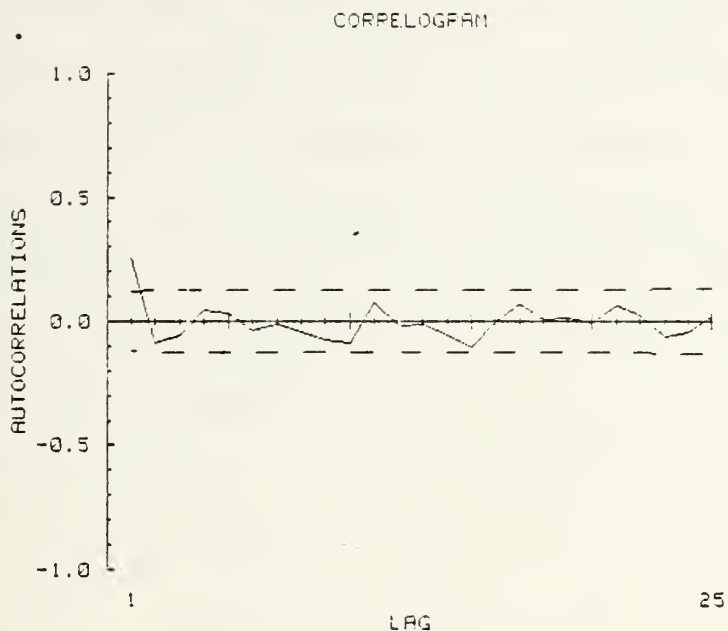


Figure 8. Correlogram of the monthly rainfall anomalies for data set RN

TABLE 3

ESTIMATED AUTOCORRELATIONS OF MONTHLY RAINFALL ANOMALIES FOR DATA SET RN

| LAG | VALUE | LAG | VALUE |
|-----|-------|-----|-------|
| 1   | .249  | 14  | -.057 |
| 2   | -.090 | 15  | -.109 |
| 3   | -.059 | 16  | -.006 |
| 4   | .041  | 17  | .067  |
| 5   | .032  | 18  | .004  |
| 6   | -.035 | 19  | .013  |
| 7   | -.011 | 20  | -.007 |
| 8   | -.043 | 21  | .063  |
| 9   | -.073 | 22  | .023  |
| 10  | -.091 | 23  | -.066 |
| 11  | .076  | 24  | -.044 |
| 12  | -.020 | 25  | .021  |
| 13  | -.012 |     |       |





### 3. Logged and Swept Data

The data should now be stationary in the means. However, as seen in Table 2, the variances of monthly rainfall amounts are not homogeneous. Kilmartin [Ref. 10] discusses various transformations of the data to remove this heteroskedacity. A plot of the variance versus mean, Figure 9 below, indicates that the logarithmic transform of the data might be useful.

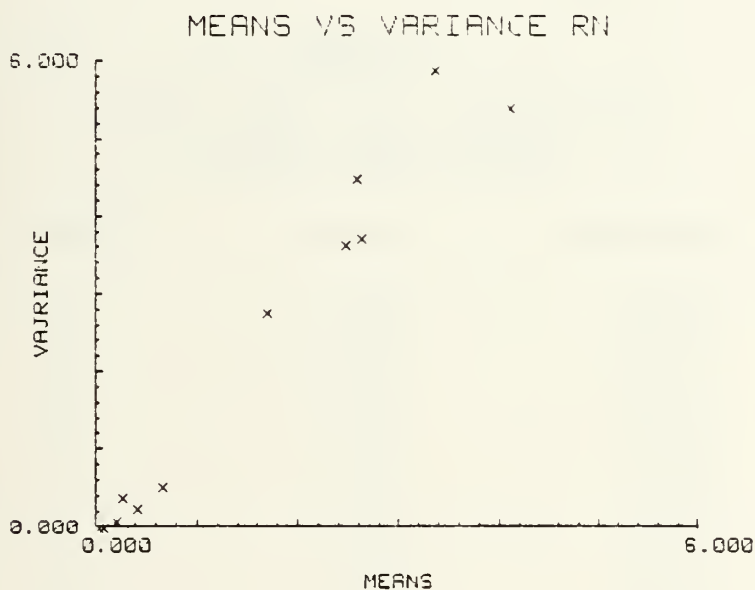


Figure 9. Plot of monthly variance against monthly means for data set RN



Since the data contain zeros, the following modified logarithmic transformation is done

$$R'_{t,m} = \ln(R_{t,m} + 1) - \frac{1}{N} \sum_{t=1}^N \ln(R_{t,m} + 1) \quad \text{II.8}$$

where the effect of the addition of the one is mostly to preserve the mapping of zeros into zeros. A more in depth discussion of this transformation is found in Kilmartin. The mapping is performed again as before and  $\tilde{R}'_{.m}$  and  $S'_{.m}$  are calculated in a manner similar to II.6 and shown in Table 4 and Figures 10 and 11.

TABLE 4  
MONTHLY MEANS AND VARIANCE FOR  
LOGGED DATA SET RN

| MONTH | MEAN  | VARIANCE |
|-------|-------|----------|
| 1     | .438  | .1549    |
| 2     | 1.092 | .3454    |
| 3     | 1.312 | .3538    |
| 4     | 1.539 | .2003    |
| 5     | 1.104 | .3795    |
| 6     | 1.133 | .3706    |
| 7     | .854  | .2728    |
| 8     | .319  | .0767    |
| 9     | .176  | .0401    |
| 10    | .054  | .0045    |
| 11    | .094  | .0094    |
| 12    | .185  | .0849    |



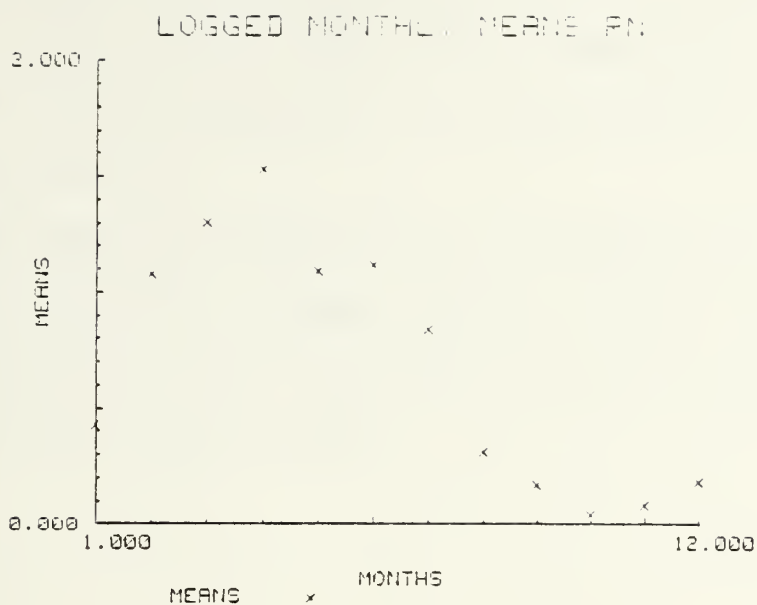


Figure 10. Monthly means of logged data set RN

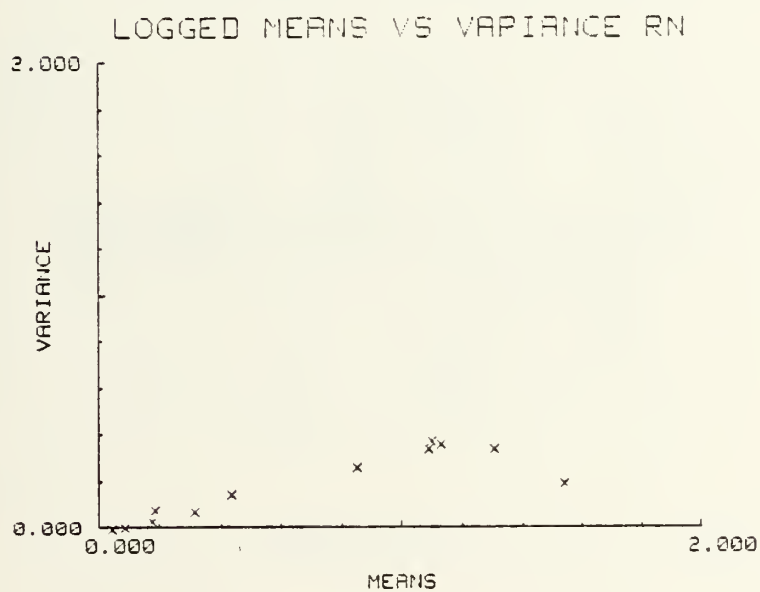


Figure 11. Plot of monthly variance against monthly means for logged data set RN



These transformations, the logarithm followed by the removal of the monthly means of the logged data, result in the series listed in Appendix A and described in Figures 12 and 13 with Table 5.

These displays indicate that a suitably stationary series has been obtained. Other methods, such as differencing, scaling, and Box-Cox transformations, see Hipel [Ref. 11], were tried but with less success.

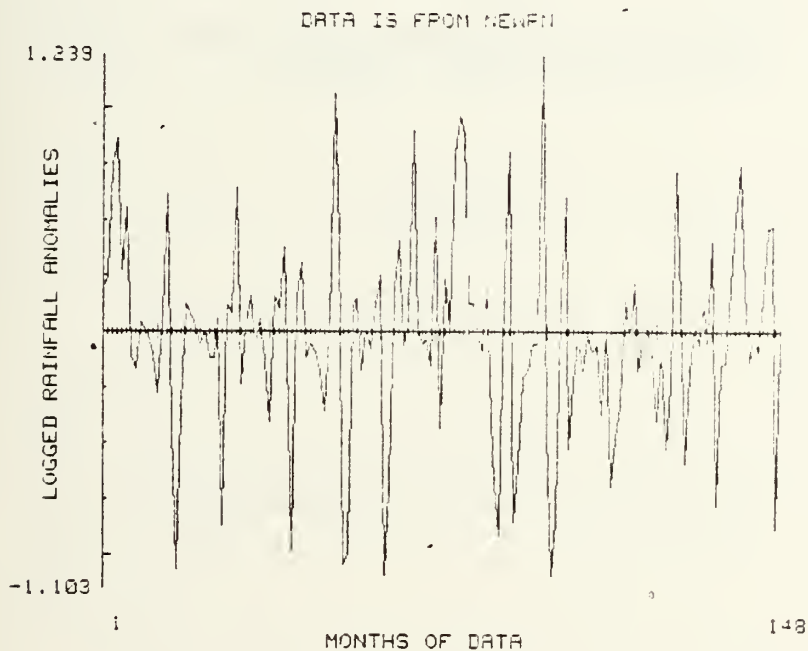


Figure 12a. Logged anomalies of monthly rainfall for data set RN. Months 1 -148





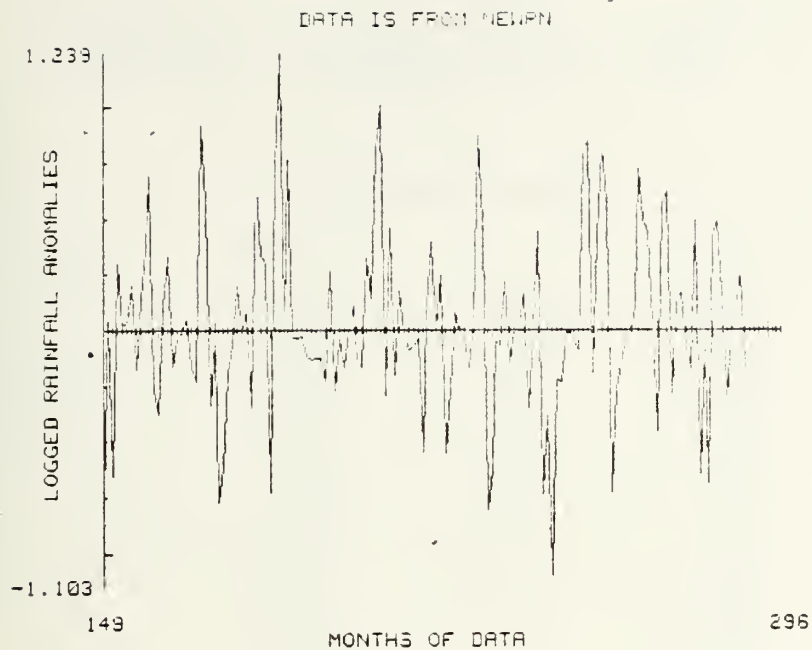


Figure 12b. Logged anomalies of monthly rainfall for data set RN

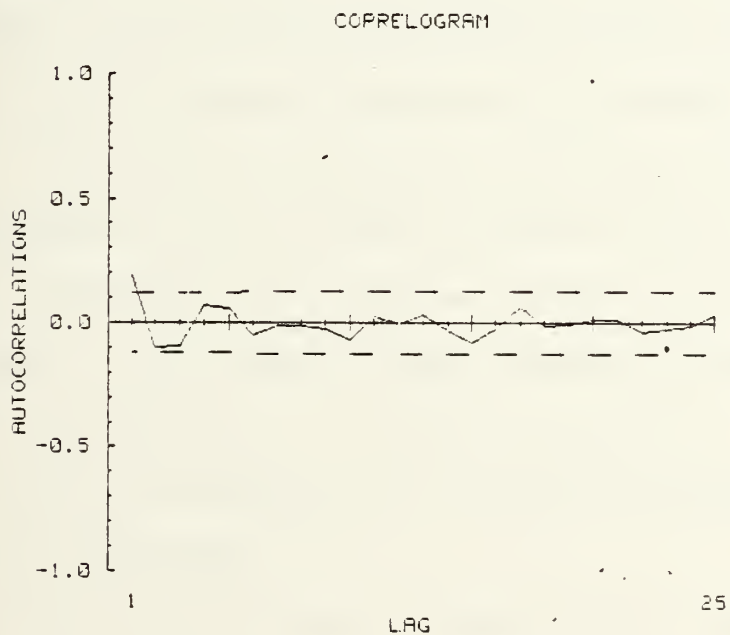


Figure 13. Correlogram of logged anomalies of monthly rainfall from data set RN



TABLE 5

ESTIMATED AUTOCORRELATIONS OF LOGGED  
ANOMALIES OF MONTHLY RAINFALL FROM  
DATA SET RN

## AUTOCORRELATIONS

| LAG | VALUE | LAG | VALUE |
|-----|-------|-----|-------|
| 1   | .191  | 14  | -.033 |
| 2   | -.102 | 15  | -.084 |
| 3   | -.095 | 16  | -.024 |
| 4   | .071  | 17  | .062  |
| 5   | .056  | 18  | -.015 |
| 6   | -.053 | 19  | -.008 |
| 7   | -.009 | 20  | .013  |
| 8   | -.014 | 21  | .013  |
| 9   | -.022 | 22  | -.038 |
| 10  | -.069 | 23  | -.024 |
| 11  | .024  | 24  | -.012 |
| 12  | -.004 | 25  | .032  |
| 13  | .032  |     |       |

## C. DATA SET FL

The label for these data derives from its location, Forest Lake, on the Monterey Peninsula, in Pebble Beach, California. Data set FL consists of monthly rainfall figures gathered by the California-American Water Company since 1896. Although this data set started quite early, the data prior to 1937 has frequent missing observations. Therefore, this data set is taken as October 1937 through September 1974, with October 1974 through September 1980 reserved for validation.

Analysis of this data set is identical to that of data set RN, therefore only the pertinent figures and tables are shown.



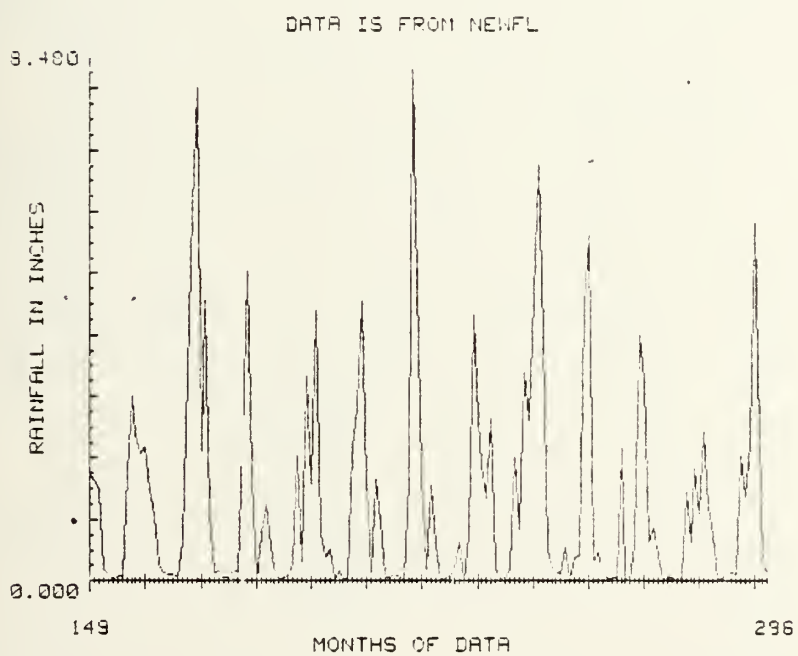
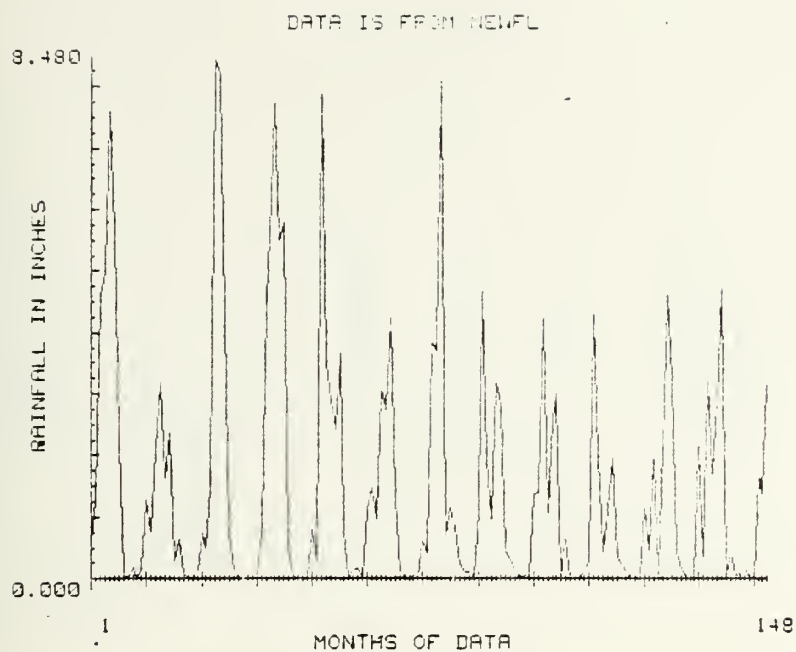


Figure 14a. Months 1 - 296 of rainfall in inches for data set FL



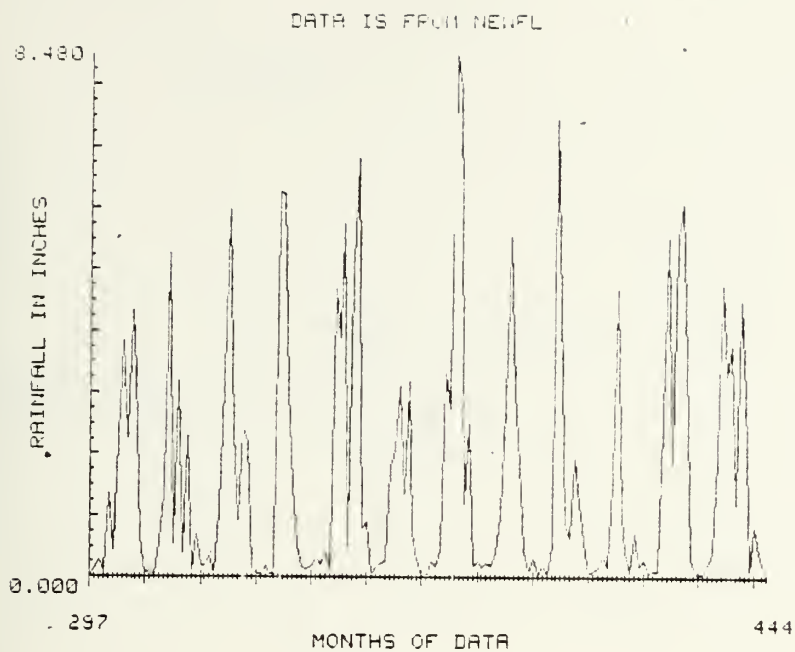


Figure 14b. Months 297-444 of rainfall in inches for data set FL

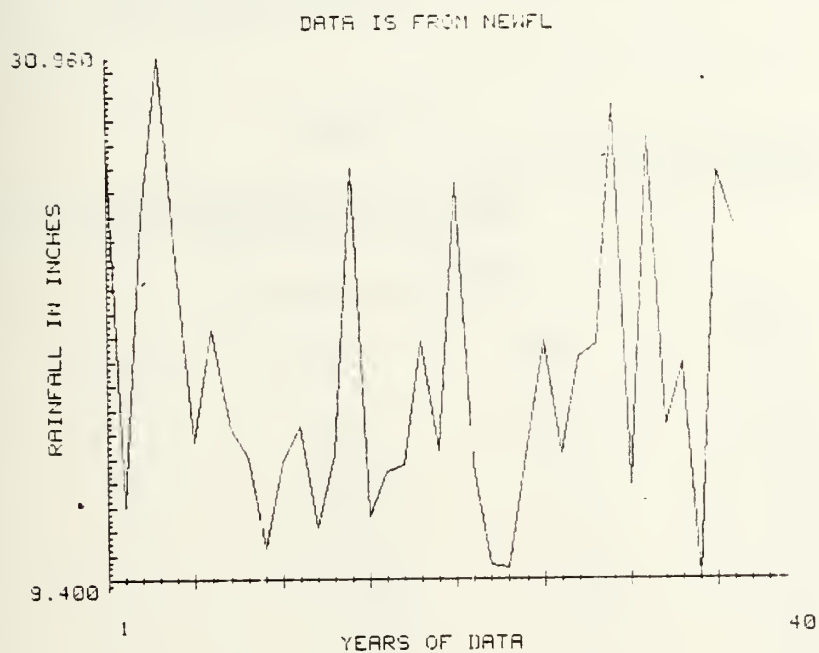


Figure 15. Yearly total rainfall for data set FL (1937 - 1974).





# CORRELOGRAM

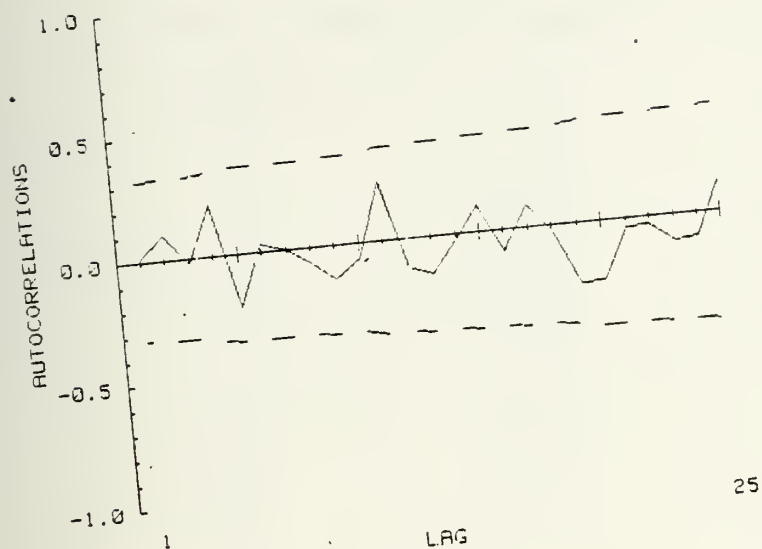


Figure 16. Correlogram of yearly total rainfall for data set FL

TABLE 6

ESTIMATED AUTOCORRELATIONS OF YEARLY TOTAL RAINFALL FOR DATA SET FL

## AUTOCORRELATIONS

| LAG | VALUE | LAG | VALUE |
|-----|-------|-----|-------|
| 1   | .010  | 14  | -.028 |
| 2   | .099  | 15  | .105  |
| 3   | -.022 | 16  | -.087 |
| 4   | .207  | 17  | .085  |
| 5   | -.214 | 18  | -.018 |
| 6   | .028  | 19  | -.245 |
| 7   | -.003 | 20  | -.237 |
| 8   | -.061 | 21  | -.035 |
| 9   | -.138 | 22  | -.031 |
| 10  | -.060 | 23  | -.106 |
| 11  | .236  | 24  | -.096 |
| 12  | -.119 | 25  | .120  |
| 13  | -.151 |     |       |



## 2. Swept Data

TABLE 7

MONTHLY MEANS AND VARIANCE FOR DATA SET FL

| MONTH | MEAN  | VARIANCE |
|-------|-------|----------|
| 1     | .744  | .4895    |
| 2     | 2.235 | 3.7444   |
| 3     | 3.049 | 4.2480   |
| 4     | 3.537 | 4.4240   |
| 5     | 2.999 | 5.9492   |
| 6     | 2.724 | 3.1743   |
| 7     | 1.559 | 2.4505   |
| 8     | .449  | .2033    |
| 9     | .153  | .0417    |
| 10    | .077  | .0081    |
| 11    | .115  | .0081    |
| 12    | .189  | .1350    |

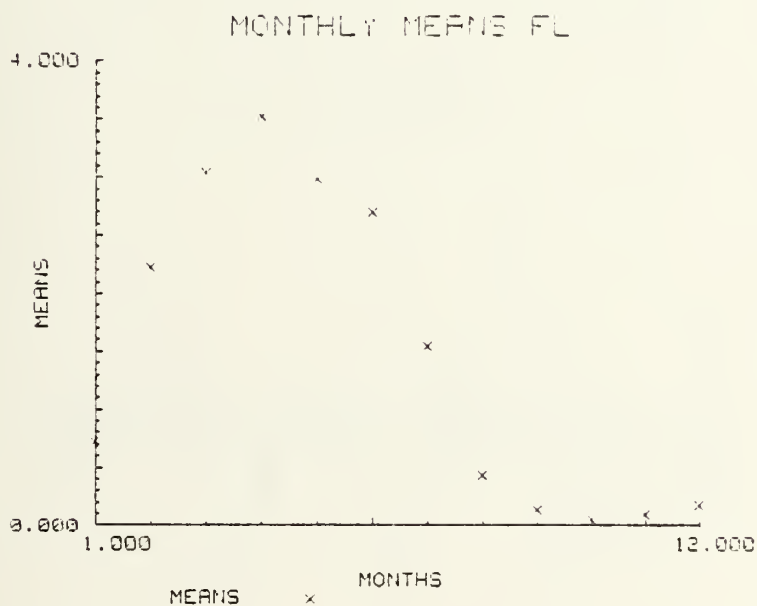


Figure 17. Monthly means for data set FL



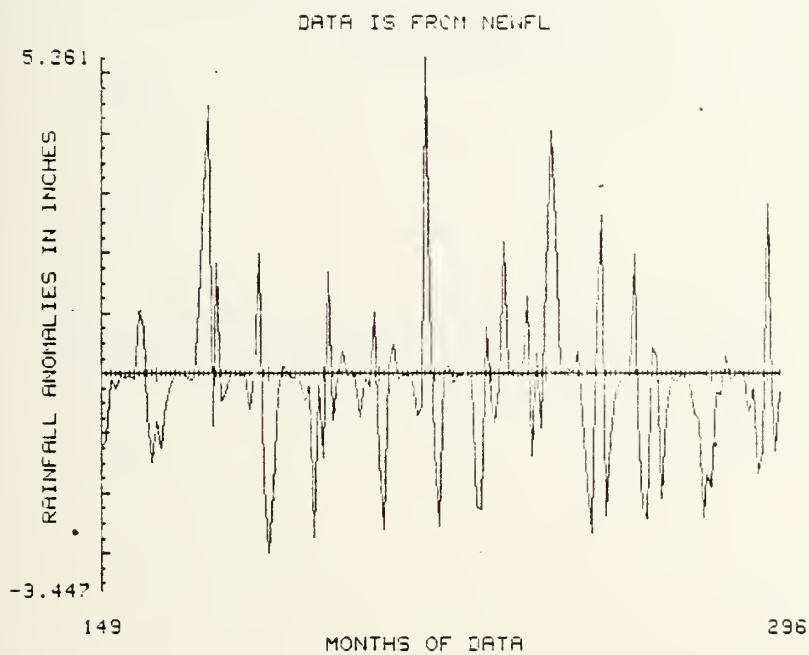
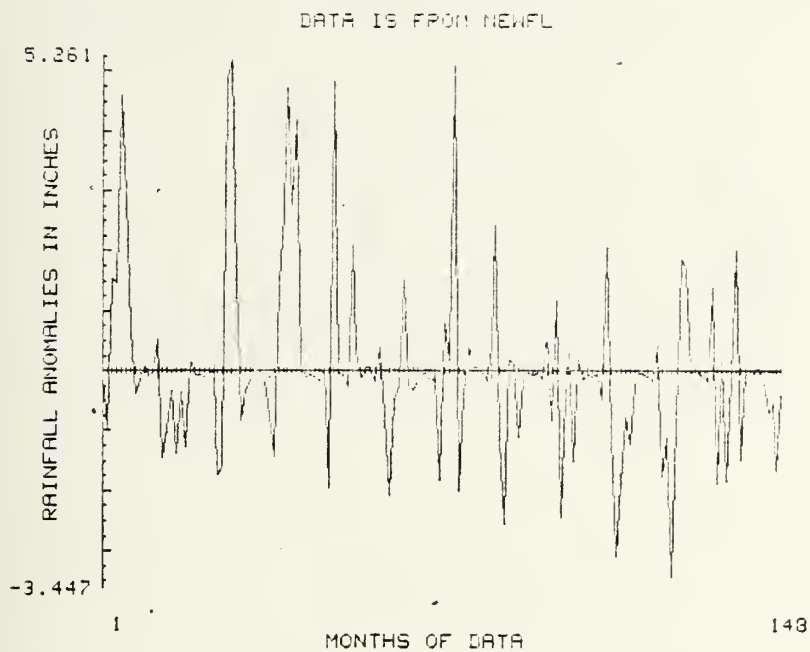


Figure 18a. Months 1 - 296 of rainfall anomalies in inches for data set FL



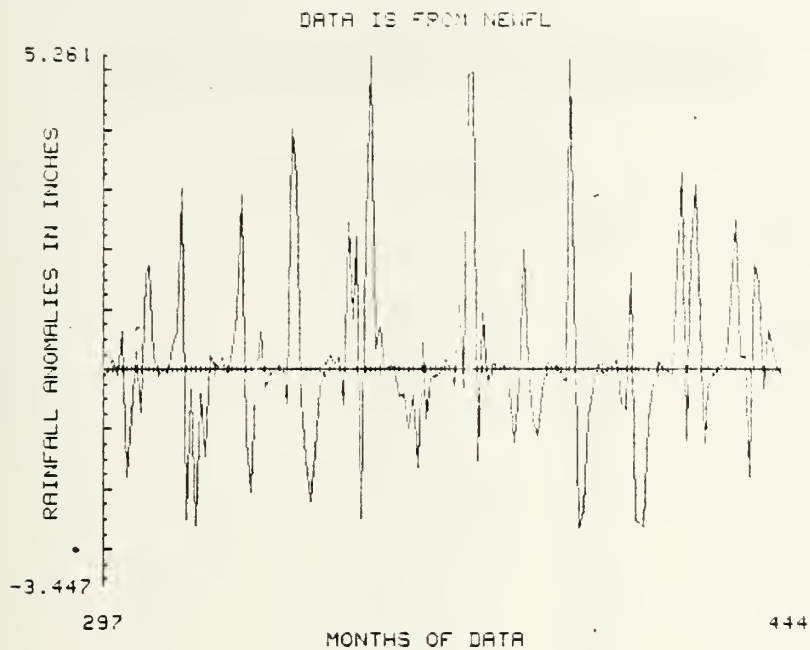


Figure 18b. Months 297-444 of rainfall anomalies in inches for data set FL

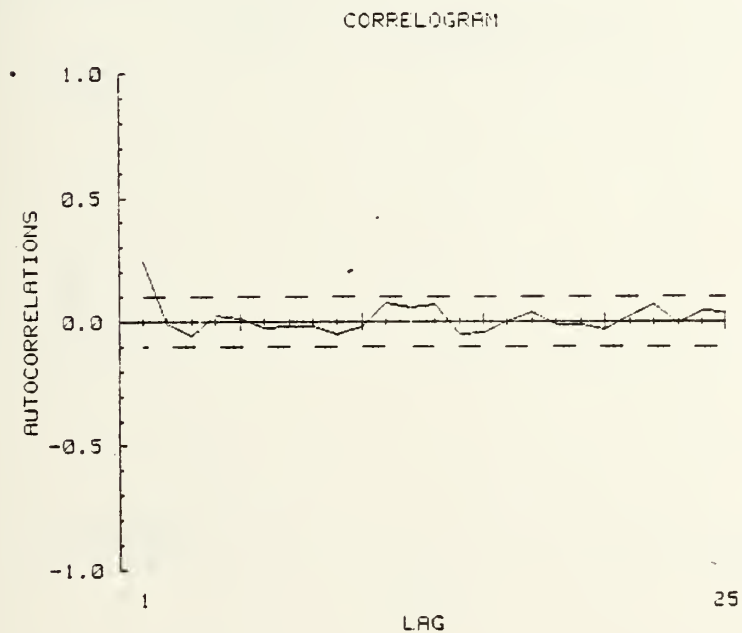


Figure 19. Correlogram of monthly rainfall anomalies for data set FL





TABLE 8

ESTIMATED AUTOCORRELATIONS OF MONTHLY  
RAINFALL ANOMALIES FOR DATA SET FL

## AUTOCORRELATIONS

| LAG | VALUE | LAG | VALUE |
|-----|-------|-----|-------|
| 1   | .244  | 14  | -.053 |
| 2   | -.007 | 15  | -.043 |
| 3   | -.056 | 16  | -.002 |
| 4   | .027  | 17  | .039  |
| 5   | .016  | 18  | -.014 |
| 6   | -.026 | 19  | -.011 |
| 7   | -.022 | 20  | -.031 |
| 8   | -.021 | 21  | .023  |
| 9   | -.051 | 22  | .067  |
| 10  | -.020 | 23  | -.007 |
| 11  | .077  | 24  | .041  |
| 12  | .059  | 25  | .039  |
| 13  | .068  |     |       |

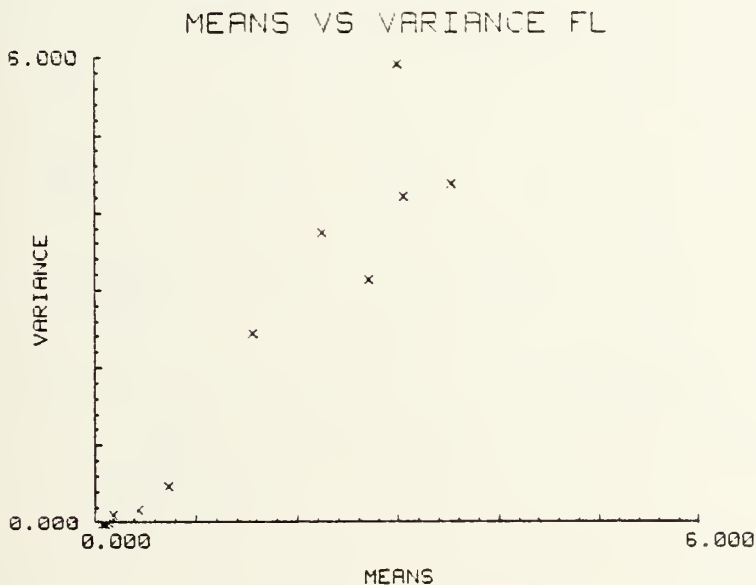
3. Logged and Swept Data

Figure 20. Plot of monthly variance against monthly means for data set FL



TABLE 9

## MONTHLY MEANS AND VARIANCE FOR LOGGED DATA SET FL

| MONTH | MEAN  | VARIANCE |
|-------|-------|----------|
| 1     | .484  | .1440    |
| 2     | 1.007 | .3440    |
| 3     | 1.269 | .2781    |
| 4     | 1.398 | .2579    |
| 5     | 1.218 | .3421    |
| 6     | 1.184 | .3006    |
| 7     | .796  | .2717    |
| 8     | .338  | .0589    |
| 9     | .130  | .0234    |
| 10    | .071  | .0061    |
| 11    | .106  | .0064    |
| 12    | .146  | .0044    |

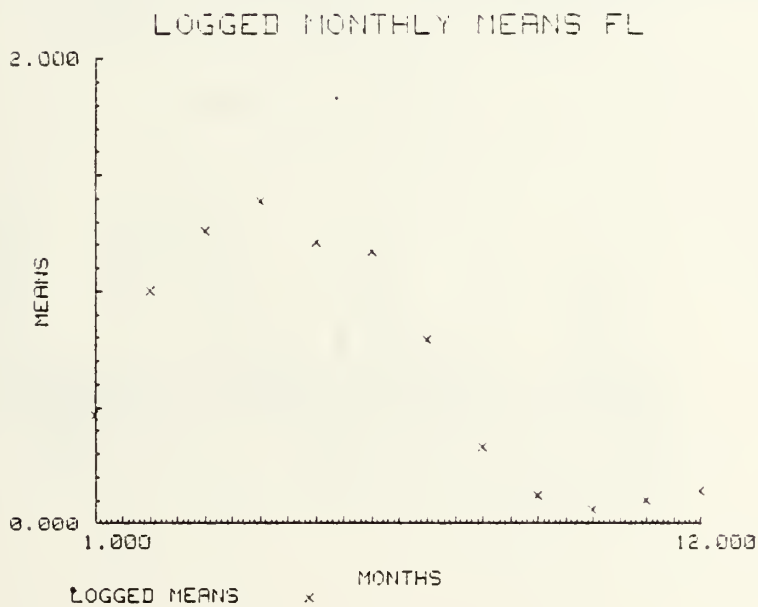


Figure 21. Monthly means of logged data set FL



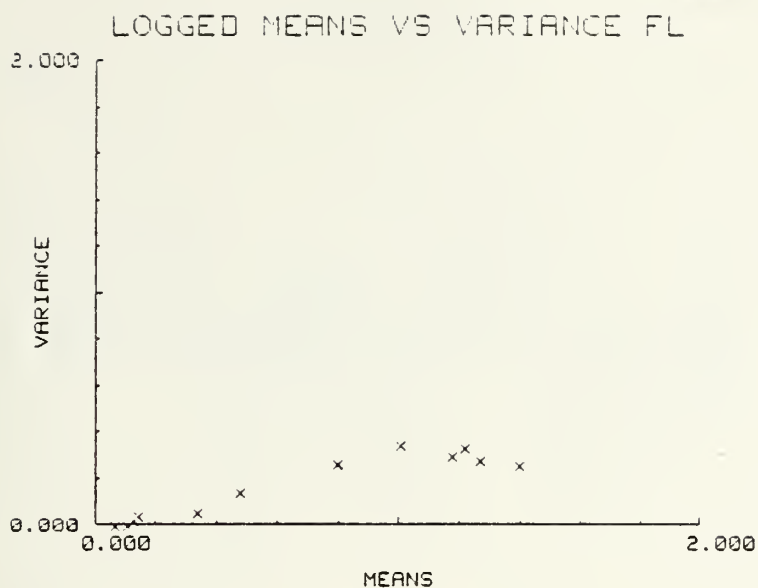


Figure 22. Plot of monthly variance against monthly means for logged data set FL

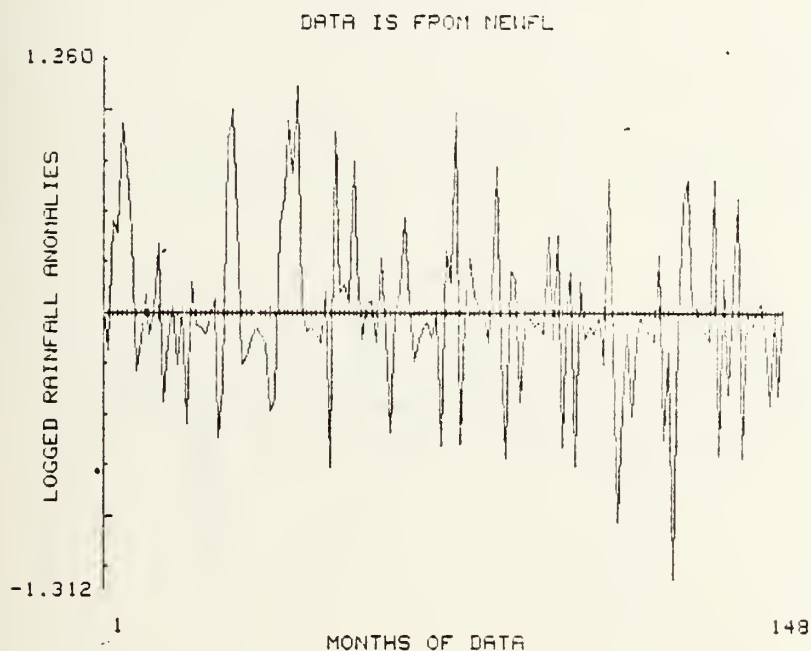


Figure 23a. Months 1 - 148 of logged rainfall anomalies for data set FL



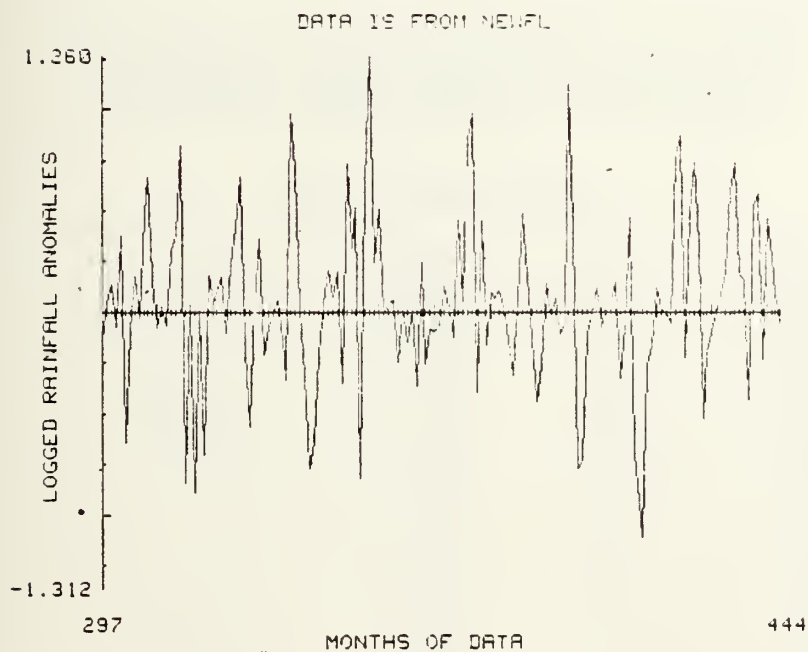
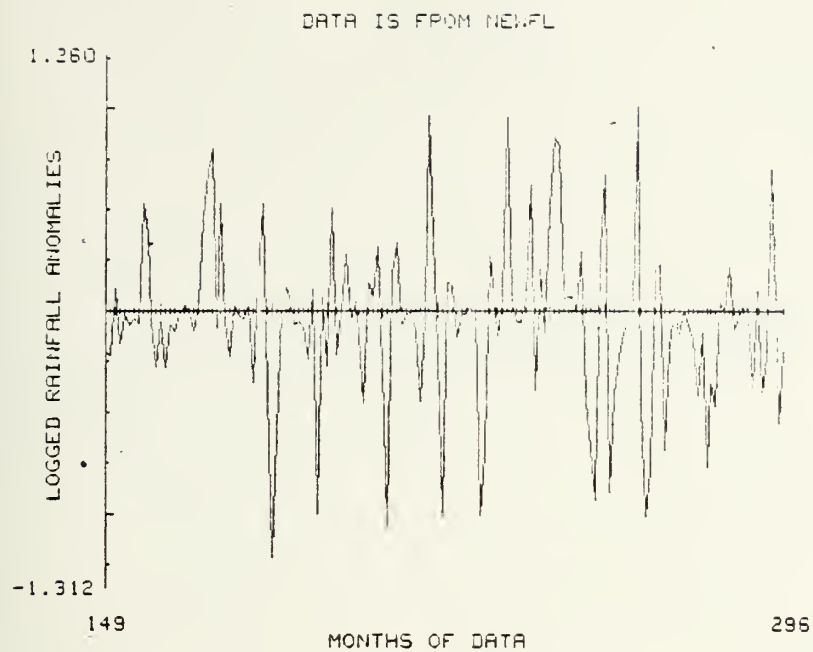


Figure 23b. Months 149 - 444 of logged rainfall anomalies from data set FL





## CORRELOGRAM

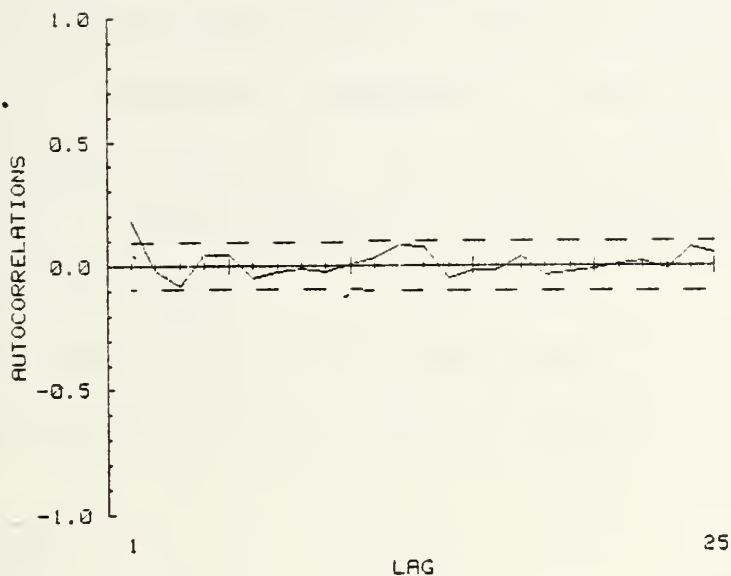


Figure 24. Correlogram of logged anomalies of monthly rainfall from data set FL.

TABLE 10

ESTIMATED AUTOCORRELATIONS OF LOGGED  
ANOMALIES FROM MONTHLY RAINFALL OF  
DATA SET FL

## AUTOCORRELATIONS

| LAG | VALUE | LAG | VALUE |
|-----|-------|-----|-------|
| 1   | .185  | 14  | -.052 |
| 2   | -.020 | 15  | -.021 |
| 3   | -.081 | 16  | -.020 |
| 4   | .046  | 17  | .040  |
| 5   | .043  | 18  | -.040 |
| 6   | -.050 | 19  | -.025 |
| 7   | -.024 | 20  | -.014 |
| 8   | -.010 | 21  | .007  |
| 9   | -.027 | 22  | .019  |
| 10  | .004  | 23  | -.015 |
| 11  | .031  | 24  | .076  |
| 12  | .084  | 25  | .047  |
| 13  | .074  |     |       |



#### D. DATA SET SC

The label for this data derives for its location, San Clemente Dam, on the Carmel River in Central California, approximately 26 kilometers southeast of data sets RN and FL on the Monterey Peninsula. Data set SC consists of monthly rainfall figures gathered by the California-American Water Company since 1926.

Analysis of this data set is again very close to that of the previous data sets and only the displays will be given.

##### 1. Raw Data

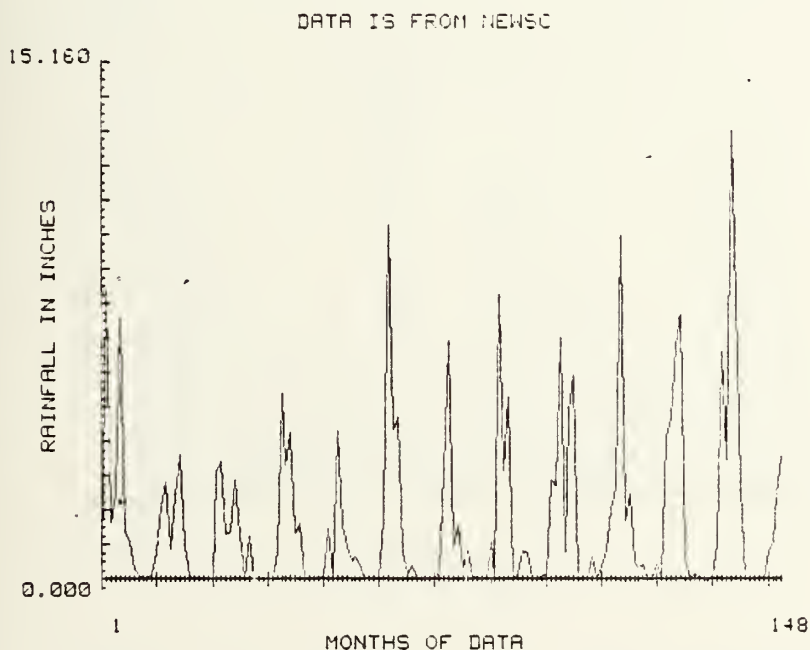


Figure 25a. Months 1 (October 1926) - 148 (January 1938) of rainfall in inches for data set SC.



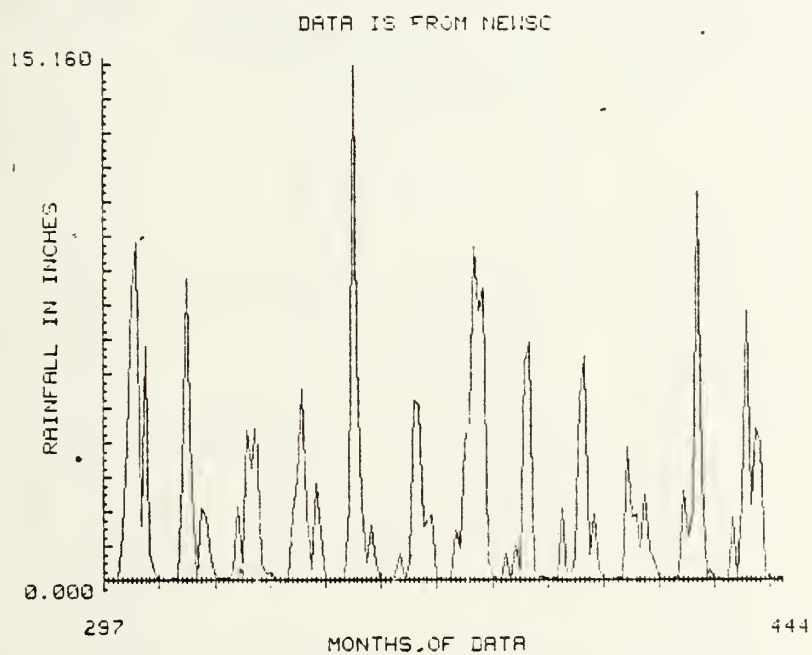
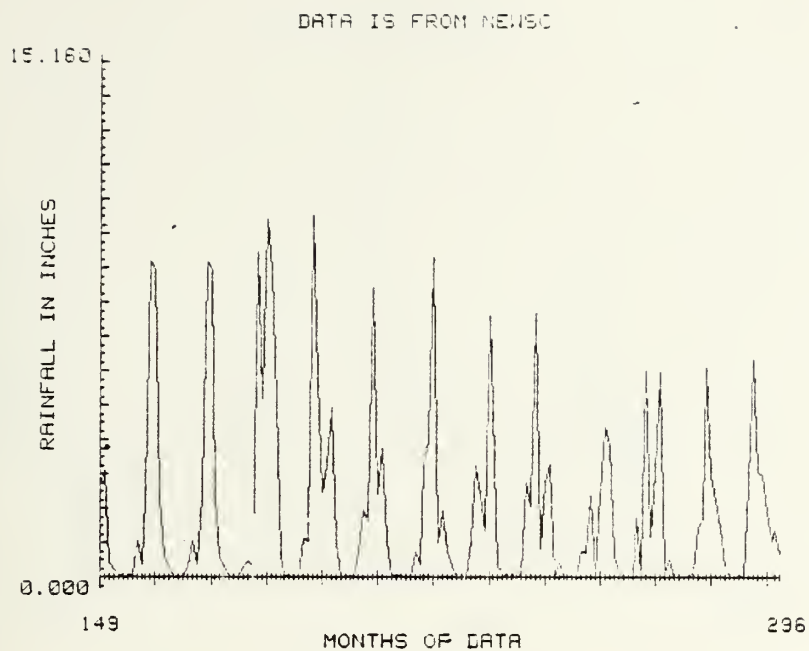


Figure 25b. Months 149 - 444 of rainfall  
for data set SC



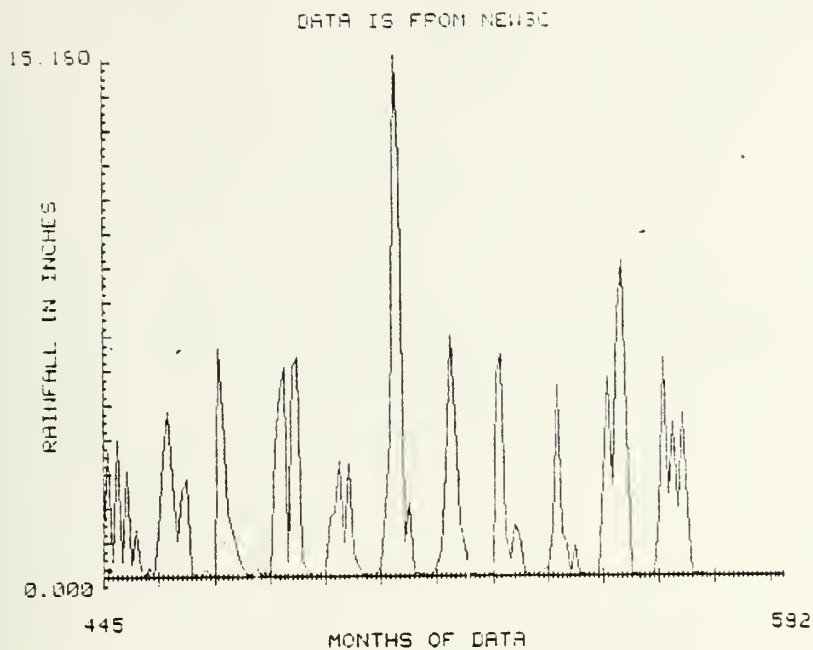


Figure 25c. Months 445 - 576 of rainfall for data set SC

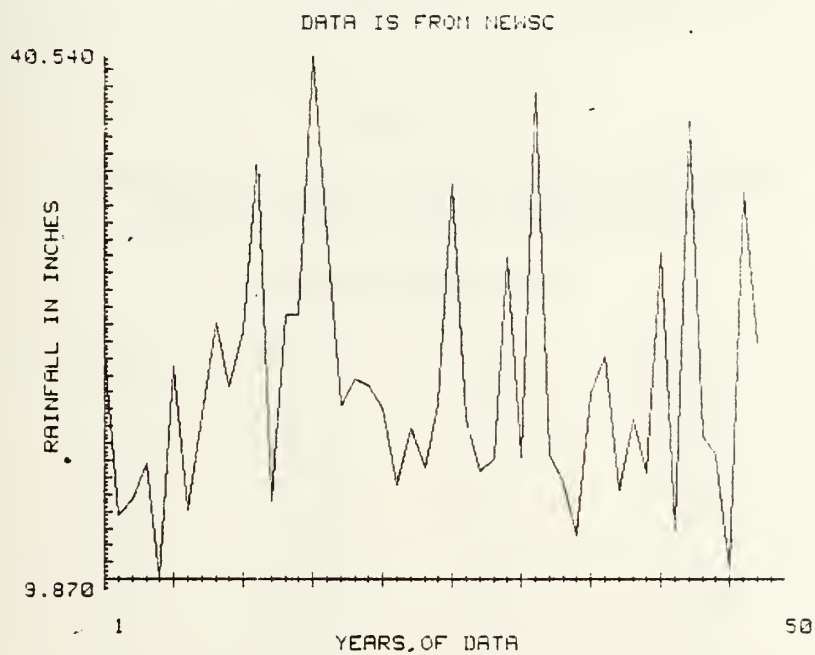


Figure 26. Yearly total rainfall for data set SC (1926 - 1974)





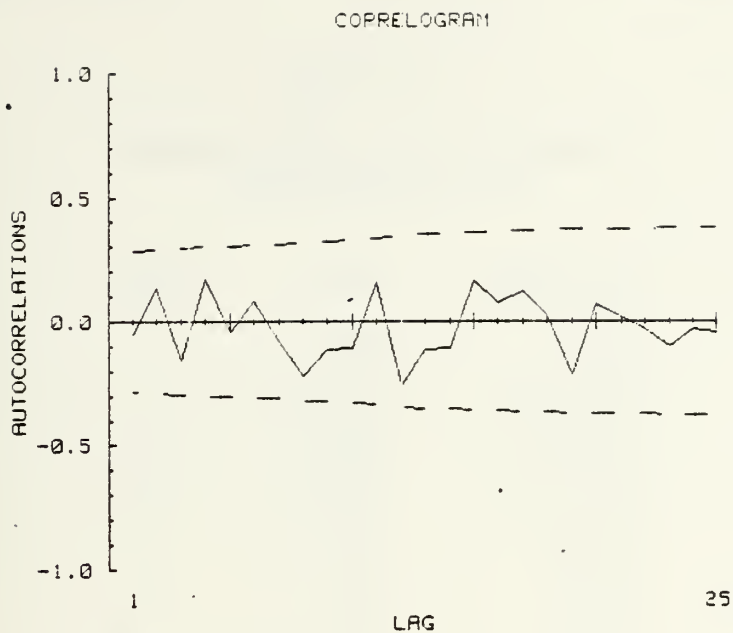


Figure 27. Correlogram of yearly total rainfall for data set SC

TABLE 11

ESTIMATED AUTOCORRELATIONS OF YEARLY  
TOTAL RAINFALL FOR DATA SET SC

AUTOCORRELATIONS

| LAG | VALUE | LAG | VALUE |
|-----|-------|-----|-------|
| 1   | -.050 | 14  | -.109 |
| 2   | .135  | 15  | .161  |
| 3   | -.158 | 16  | .077  |
| 4   | .168  | 17  | .116  |
| 5   | -.042 | 18  | .025  |
| 6   | .081  | 19  | -.214 |
| 7   | -.084 | 20  | .066  |
| 8   | -.217 | 21  | .019  |
| 9   | -.111 | 22  | -.028 |
| 10  | -.107 | 23  | -.101 |
| 11  | .158  | 24  | -.030 |
| 12  | -.260 | 25  | -.042 |
| 13  | -.110 |     |       |



## 2. Swept Data

TABLE 12

MONTHLY MEANS AND VARIANCES  
FOR DATA SET SC

| MONTH | MEAN  | VARIANCE |
|-------|-------|----------|
| 1     | .698  | .5945    |
| 2     | 2.175 | 4.2382   |
| 3     | 3.940 | 10.1783  |
| 4     | 4.599 | 8.4899   |
| 5     | 4.353 | 13.3443  |
| 6     | 3.080 | 5.2744   |
| 7     | 1.700 | 3.4486   |
| 8     | .431  | .1912    |
| 9     | .111  | .0451    |
| 10    | .017  | .0055    |
| 11    | .037  | .0125    |
| 12    | .103  | .1040    |

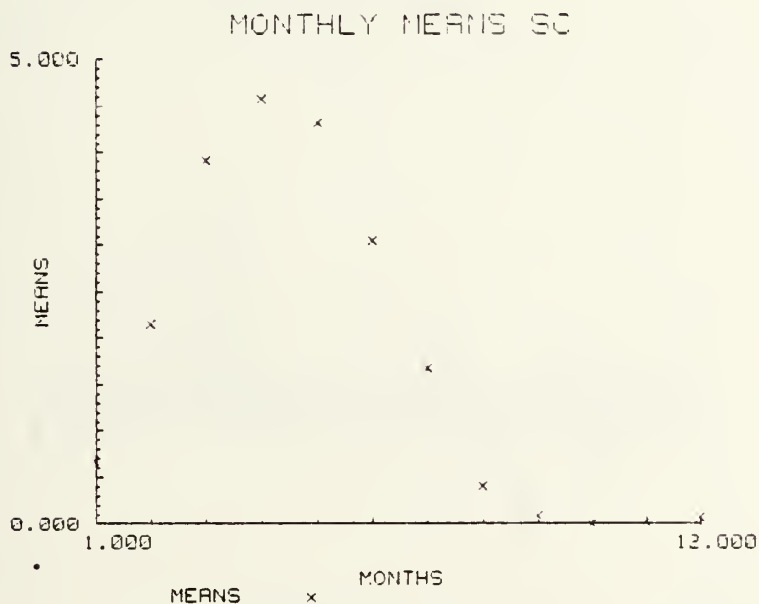
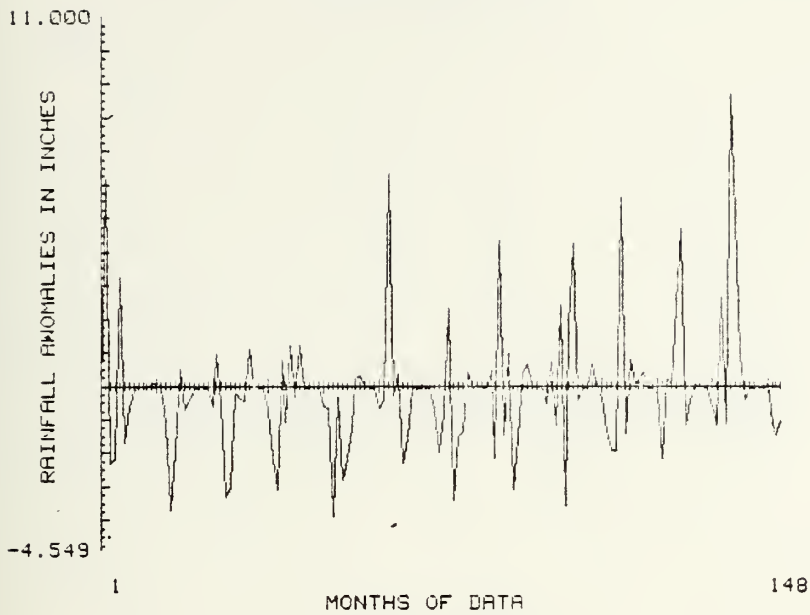


Figure 28. Monthly means for data set SC



DATA IS FROM NEWSO -



DATA IS FROM NEWSO

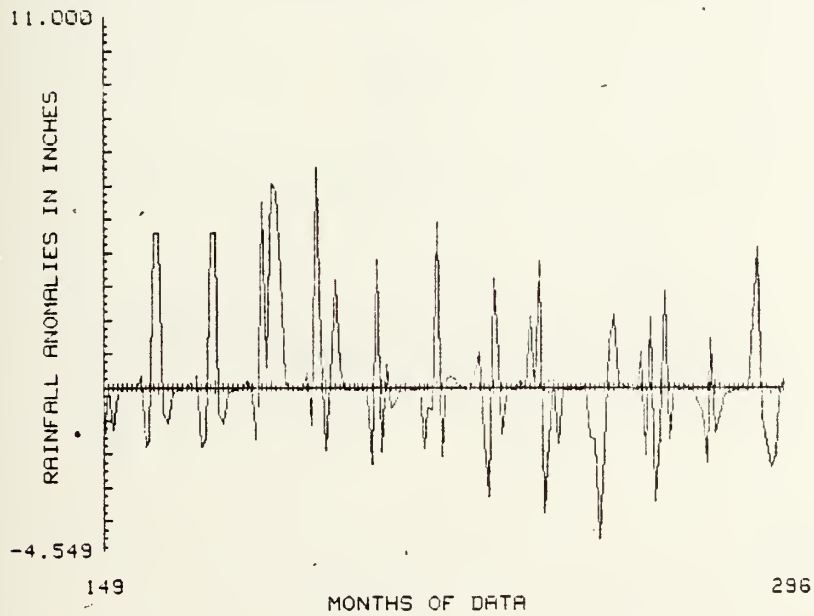


Figure 29a. Months 1 - 296 anomalies in inches for data set SC



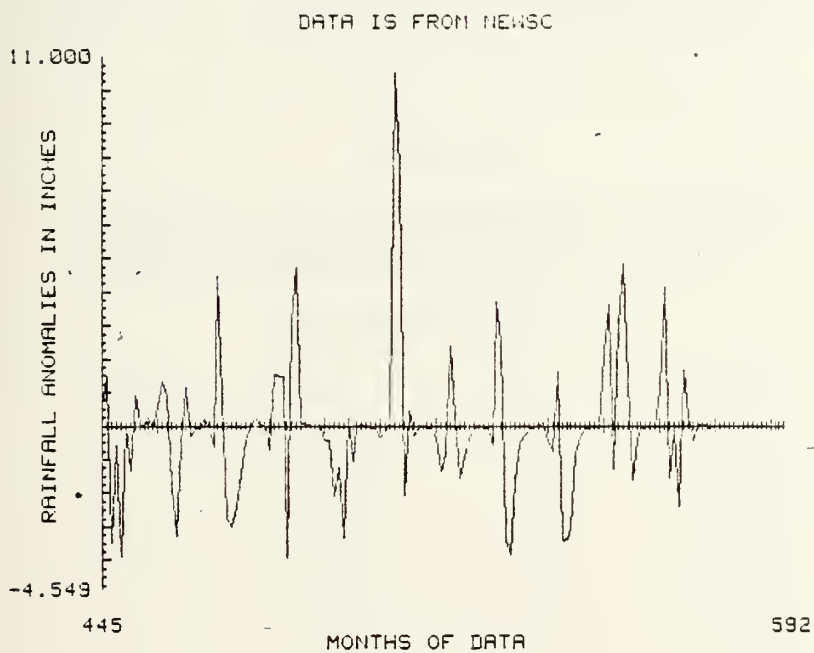
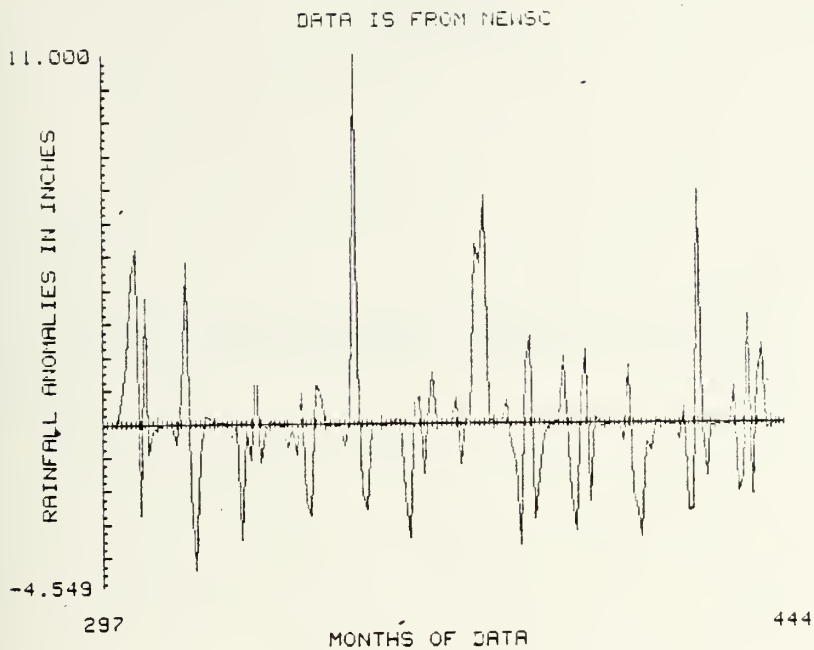


Figure 29b. Months 297 - 576 anomalies in inches for data set SC





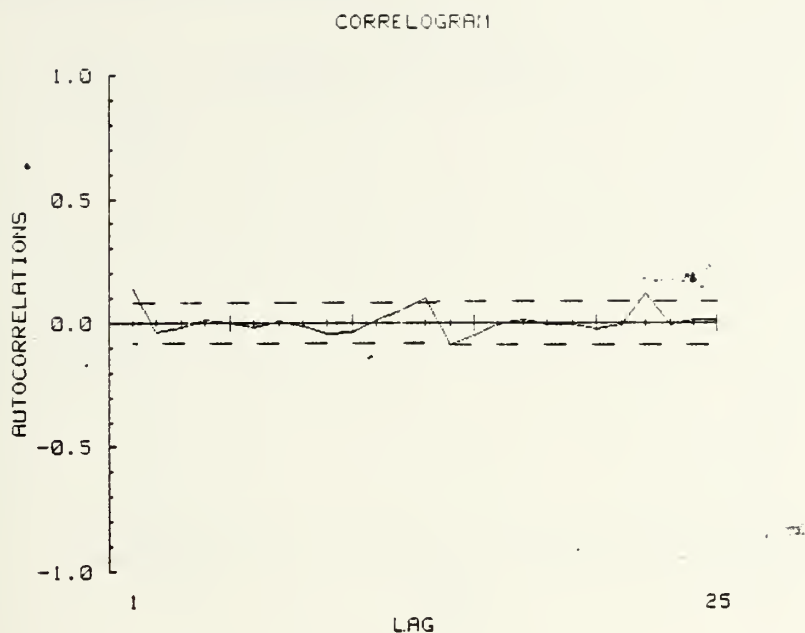


Figure 30. Correlogram of monthly rainfall anomalies for data set SC

TABLE 13

ESTIMATED AUTOCORRELATIONS OF MONTHLY  
RAINFALL ANOMALIES FOR DATA SET SC

| AUTOCORRELATIONS |       |     |       |
|------------------|-------|-----|-------|
| LAG              | VALUE | LAG | VALUE |
| 1                | .140  | 14  | -.088 |
| 2                | -.039 | 15  | -.051 |
| 3                | -.021 | 16  | -.006 |
| 4                | .012  | 17  | .015  |
| 5                | -.001 | 18  | -.006 |
| 6                | -.019 | 19  | -.008 |
| 7                | .003  | 20  | -.027 |
| 8                | -.013 | 21  | -.003 |
| 9                | -.038 | 22  | -.122 |
| 10               | -.038 | 23  | -.006 |
| 11               | .014  | 24  | .011  |
| 12               | .051  | 25  | .011  |
| 13               | .102  |     |       |



### 3. Logged and Swept Data

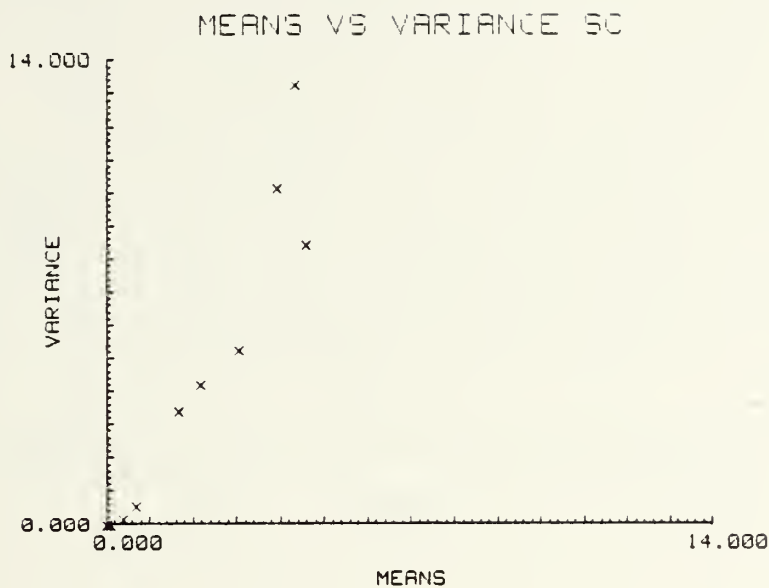


Figure 31. Plot of monthly variance against anomaly means for data set SC.

TABLE 14

MONTHLY MEANS AND VARIANCES OF  
LOGGED DATA SET SC

| MONTH | MEAN  | VARIANCE |
|-------|-------|----------|
| 1     | .444  | .1623    |
| 2     | .961  | .3999    |
| 3     | 1.408 | .3949    |
| 4     | 1.583 | .3117    |
| 5     | 1.444 | .4928    |
| 6     | 1.243 | .3556    |
| 7     | .817  | .3247    |
| 8     | .320  | .0726    |
| 9     | .092  | .0227    |
| 10    | .015  | .0040    |
| 11    | .031  | .0083    |
| 12    | .075  | .0351    |



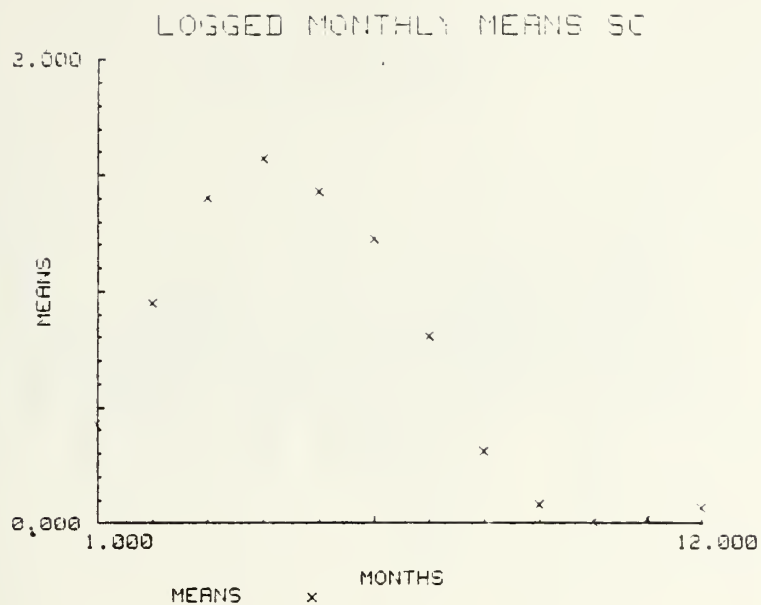


Figure 32. Monthly means of logged data set SC

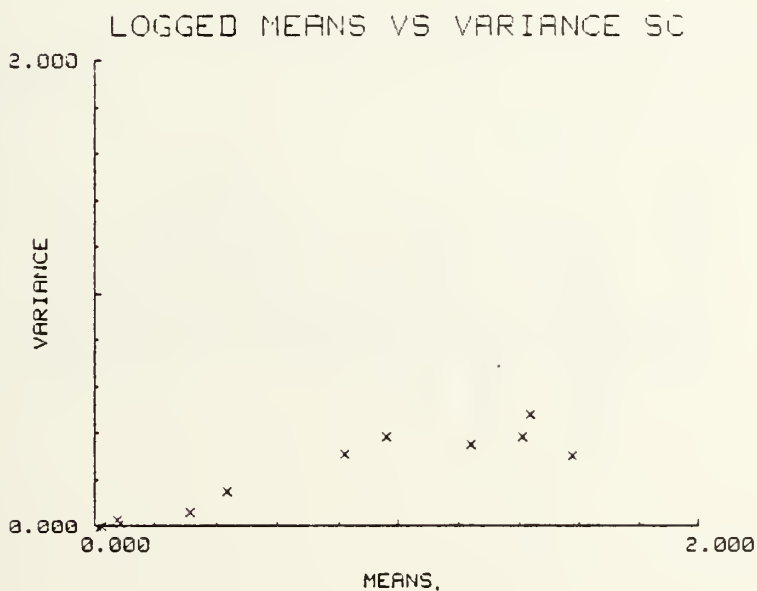


Figure 33. Plot of monthly variance against monthly means for logged data set SC



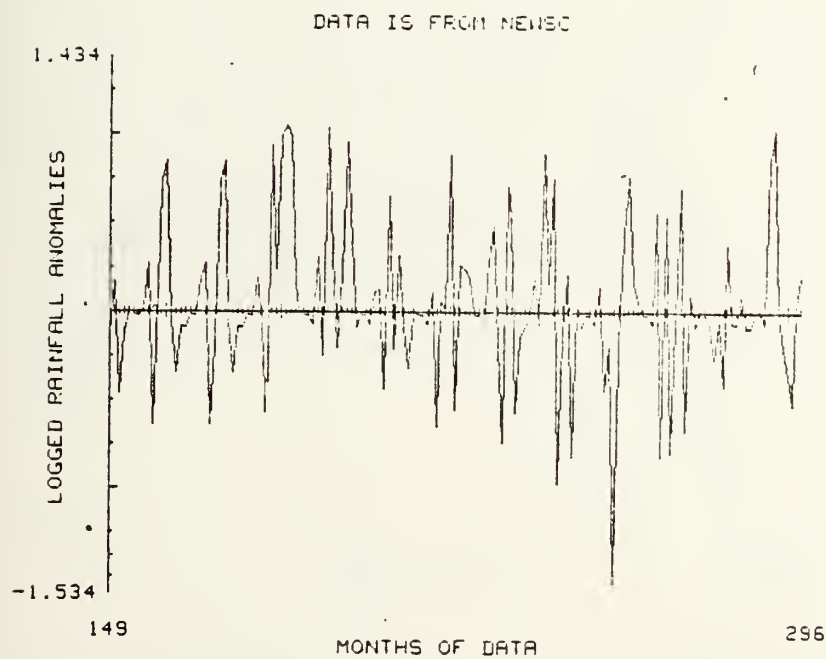
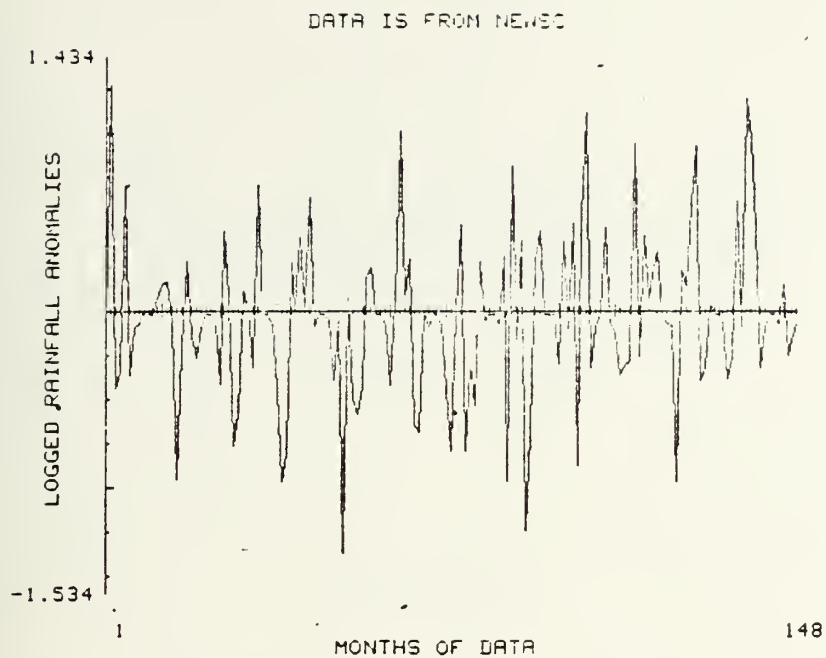


Figure 34a. Months 1 - 296 of logged rainfall anomalies from data set SC





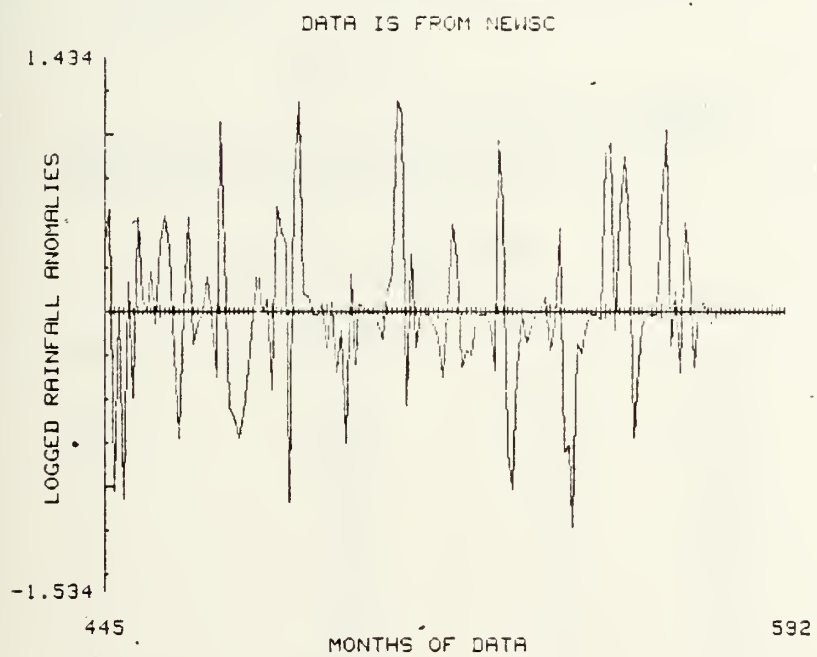
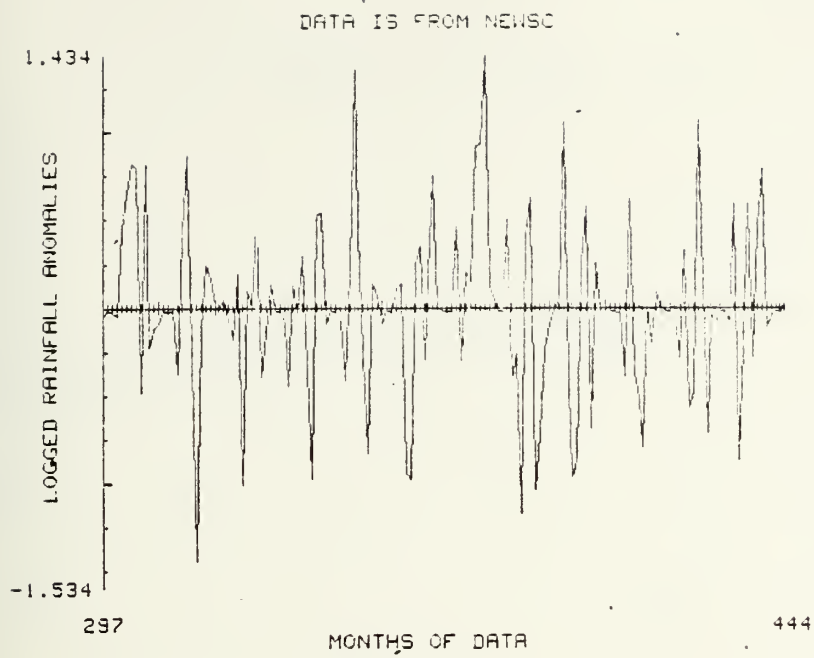


Figure 34b. Months 297 - 576 of logged rainfall anomalies from data set SC



# CORRELOGRAM

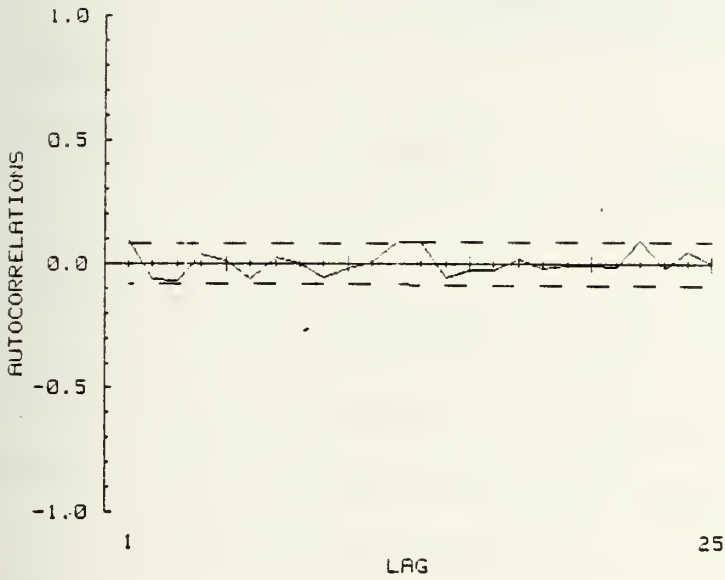


Figure 35. Correlogram of logged anomalies of monthly rainfall from data set SC

TABLE 15

ESTIMATED AUTOCORRELATION OF LOGGED  
ANOMALIES OF MONTHLY RAINFALL FROM  
DATA SET SC

## AUTOCORRELATIONS

| LAG | VALUE | LAG | VALUE |
|-----|-------|-----|-------|
| 1   | .096  | 14  | -.057 |
| 2   | -.065 | 15  | -.022 |
| 3   | -.066 | 16  | -.023 |
| 4   | .038  | 17  | .021  |
| 5   | .012  | 18  | -.017 |
| 6   | -.061 | 19  | -.005 |
| 7   | .023  | 20  | -.008 |
| 8   | -.001 | 21  | -.011 |
| 9   | -.056 | 22  | .092  |
| 10  | -.021 | 23  | -.019 |
| 11  | .007  | 24  | .050  |
| 12  | .091  | 25  | .002  |
| 13  | .091  |     |       |



### III. FIRST ORDER MARKOV MODEL

#### A. THEORY

As first shown by equation I.1, the general ARMA(p,q) model is:

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad \text{III.1}$$

The development and discussion of this type of model is contained in detail in Box and Jenkin [Ref. 1] and Nelson [Ref. 8]. The modeling process is a three fold procedure. The parts are:

- (1) Identification
- (2) Estimation
- (3) Diagnosis.

Identification is conducted using the correlogram and a plot of the partial-autocorrelations (or partial correlogram). The partial autocorrelations are related to the autocorrelations, see Box and Jenkins [Ref. 1], Nelson [Ref. 8], or Richards and Woodall [Ref. 12]. These partial autocorrelations are used to determine the order of the moving average process much like the autocorrelations may be used to determine the order of the auto-regressive process.

Once the autocorrelations and partial autocorrelations have been found, the degree of the ARMA may be estimated by



techniques described in Box and Jenkins, Nelson or Richards and Woodall. Each of the data sets, once logged and swept, indicated that the most probable model was an ARMA(1,0) or AR(1) or more commonly a first-order autoregressive Markov model. This model is simply;

$$\tilde{z}_t = \rho \tilde{z}_{t-1} + a_t \quad \text{III.2}$$

where the  $\rho$  is the autocorrelation of lag one. Thus, this model indicates that any persistence in the data are conditionally independent of the past given the lag one value.

Subsections B, C, and D below show this model as applied to the three data sets of interest. The residuals of the model  $\tilde{z}_t - \rho \tilde{z}_{t-1}$  are examined. The residuals appear to be independent, however, they do not appear to be normally distributed; for example, there is a high peak around zero. One possible reason for this discrepancy may be the dichotomy of winter and summer rain as indicated in Tables 2, 4, 7, 9, 12, and 14. The existence of months with zero rainfall during the summer suggests that one should consider the summer, when rain is sparse, completely separate from the winter when rain is more abundant. Therefore, also shown in the subsections below is the autoregressive model applied to the winter months only. This is accomplished by stripping out months 9 through 12 (June through September) of the data sets and treating the remaining data as a continuous set. In other words, the first ten months are then

$R_{1,1}, R_{1,2}, R_{1,3}, R_{1,4}, R_{1,5}, R_{1,6}, R_{1,8}, R_{2,1}, R_{2,2}.$





The appropriate correlograms and partial correlograms are displayed prior to the model applications.

## B. DATA SET RN

### 1. Twelve Month Series

This data set is described in section II.b. The remaining diagnostic device needed is the partial correlogram of Figure 36 and the corresponding values in Table 16.

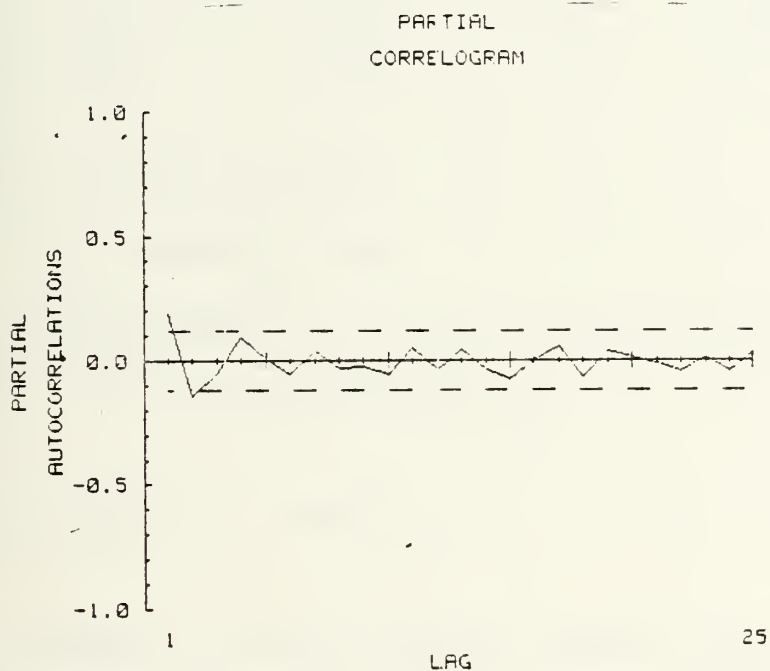


Figure 36. Partial correlogram of the logged rainfall anomalies of data set RN



TABLE 16

ESTIMATED PARTIAL-AUTOCORRELATIONS FOR  
LOGGED RAINFALL ANOMALIES OF DATA SET RN

| LAG | VALUE | LAG | VALUE |
|-----|-------|-----|-------|
| 1   | .191  | 14  | -.040 |
| 2   | -.144 | 15  | -.073 |
| 3   | -.047 | 16  | .001  |
| 4   | .092  | 17  | .054  |
| 5   | .006  | 18  | -.067 |
| 6   | -.058 | 19  | .039  |
| 7   | .037  | 20  | .012  |
| 8   | -.034 | 21  | -.001 |
| 9   | -.028 | 22  | -.043 |
| 10  | -.056 | 23  | .013  |
| 11  | .048  | 24  | -.043 |
| 12  | -.041 | 25  | -.034 |
| 13  | .047  |     |       |

The model of interest is then

$$\tilde{z}'_t = .191\tilde{z}'_{t-1} + a_t \quad \text{III.3}$$

where the random shocks  $\{a_t\}$  are assumed to be distributed iid  $N(0, \sigma_a^2)$  and  $\sigma_a^2$  is estimated as

$$\frac{1}{N-1} \sum_{t=1}^N (\tilde{z}'_t - .191\tilde{z}'_{t-1})^2 \quad \text{III.4}$$

The goodness of this fit may be viewed in two ways. Firstly, are the residuals,  $\{a_t\}$  independent? Secondly, are the residuals distributed as Normal (Gaussian) random variables? A plot of the residuals follows in Figure 37. The question of independence is addressed in Figure 38 (Correlogram), Figure 39 (Lag one plot), Figure 40 (Residuals vs. lag one), and Table 17 (Turning points). For a discussion of the usefulness of the turning points see Kendall [Ref. 14].



All of these displays and tests tend to indicate that the residuals are in fact serially independent. The statistics of the residuals are in Table 18. A Normal Plot of the residuals (Figure 41), in which the sample is normalized by removing the mean and scaling by the standard deviation and then plotted on normal paper, should yield a nearly straight line corresponding to the dashed line of the figure. The Normal Plot accompanied by the sample histogram (Figure 42) addresses the normality of these data. As may be seen from the kurtosis, the fluctuations of the sample CDF near the midpoint, and the peak of the histogram, the normality of this data are questionable. To confirm this a chi-squared goodness of fit test was conducted yielding a value of 49.18 with 17 degrees of freedom, again rejecting any hypothesis of normality at a significance level of  $5 \times 10^{-5}$ .

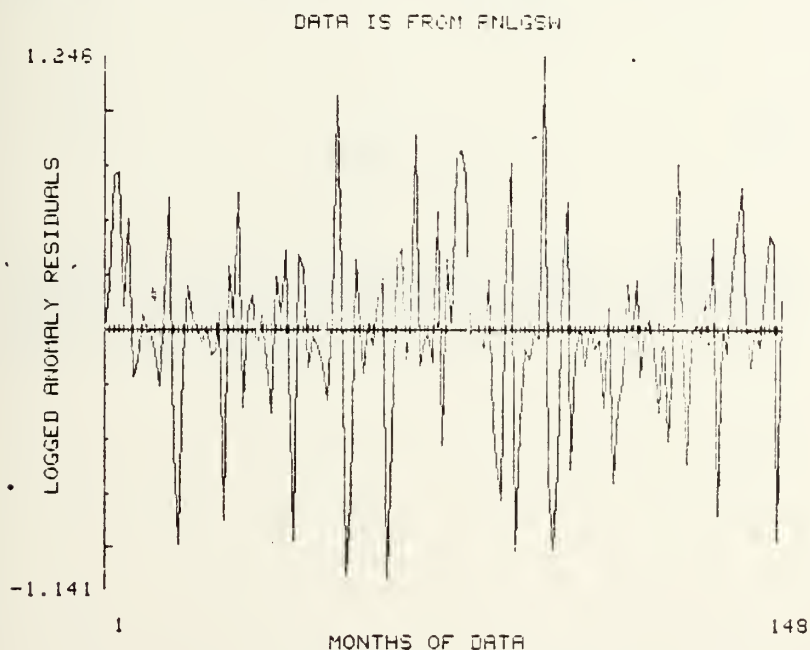


Figure 37. First order Markov residuals from logged rainfall anomalies of data set RN. Months 149 - 292



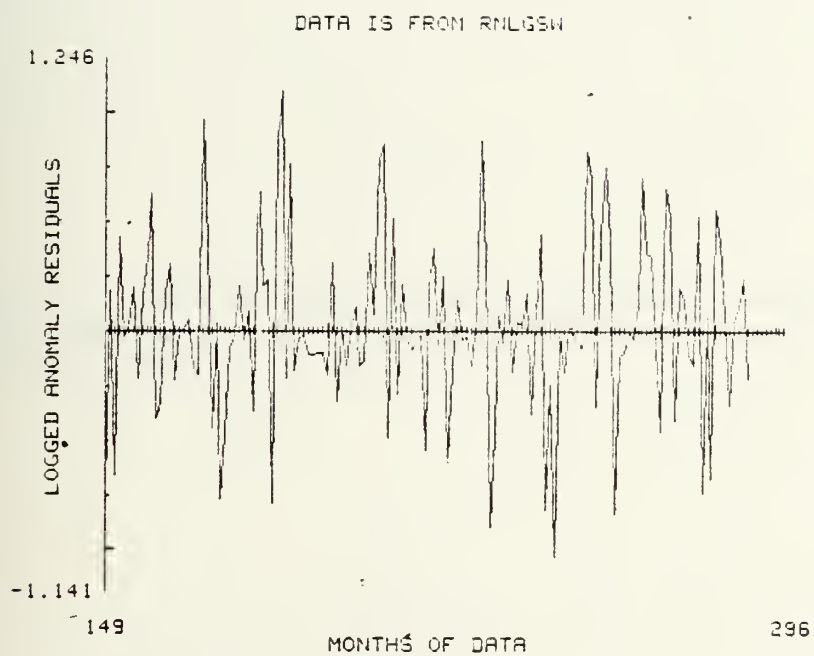


Figure 37b. First order Markov residuals from  
logged rainfall anomalies of data set RN.  
Months 149 - 292





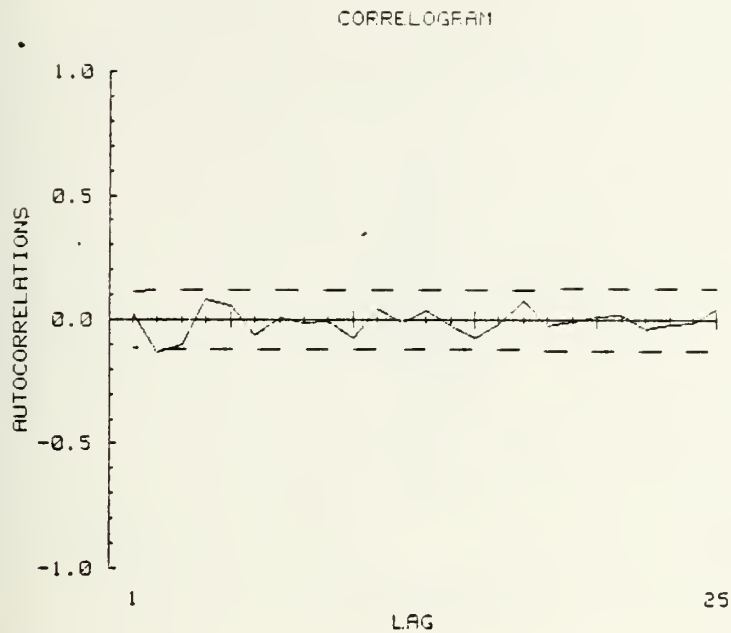


Figure 38. Auto correlations of residuals from first order Markov process applied to the logged rainfall anomalies of data set RN

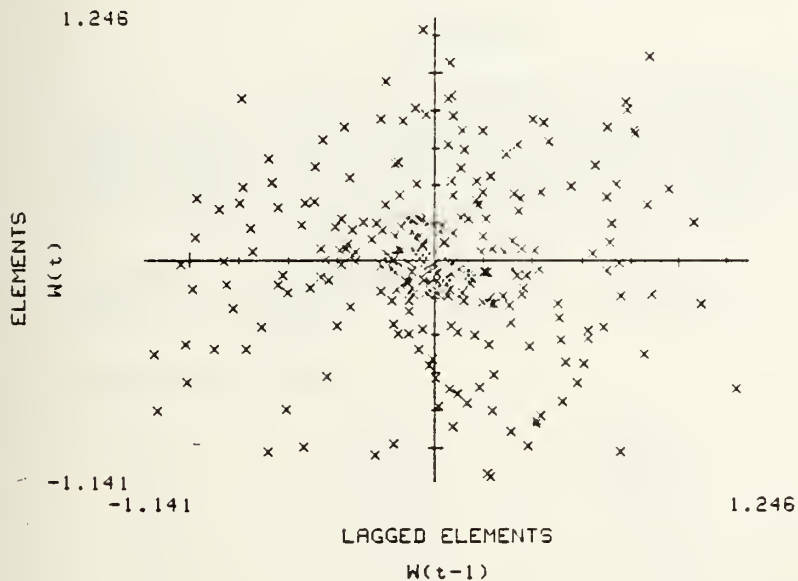


Figure 39. Lag one plot of first order Markov residuals from logged rainfall anomalies of data set RN



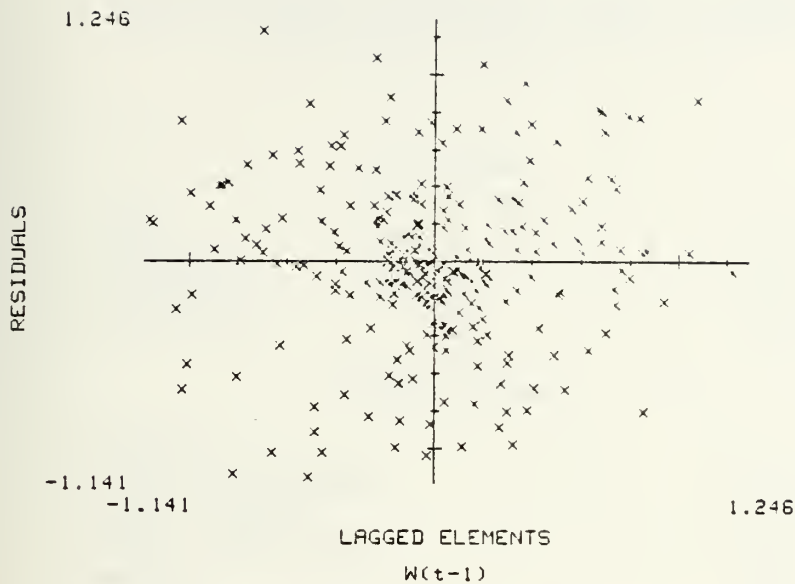


Figure 40. First order Markov residuals versus lag one data point from logged rainfall anomalies of data set RN

TABLE 17

ACTUAL AND EXPECTED NUMBER OF TURNING POINTS  
AND ACTUAL AND EXPECTED PHASE FREQUENCIES  
FOR THE FIRST ORDER MARKOV RESIDUALS FROM  
THE LOGGED RAINFALL ANOMALIES OF DATA SET RN

NUMBER OF TURNING POINTS = 191  
 $E[P] = 190.667$        $V[P] = 15.899$

PHASE LENGTHS

| D | OBS. | $E[*]$ |
|---|------|--------|
| 1 | 117  | 118.8  |
| 2 | 56   | 52.1   |
| 3 | 15   | 14.9   |
| 4 | 3    | 3.2    |
| 5 | 0    | .6     |
| 6 | 0    | .1     |
| 7 | 0    | 0.0    |



TABLE 18

GENERAL STATISTICS OF FIRST ORDER  
MARKOV RESIDUALS FROM LOGGED RAINFALL  
ANOMALIES OF DATA SET RN

Moments

|          |       |
|----------|-------|
| Mean     | -.001 |
| Variance | .117  |
| Skewness | -.066 |
| Kurtosis | .523  |

Percentiles

|                 |        |
|-----------------|--------|
| Minimum         | -1.141 |
| Lower Sixteenth | -.745  |
| Lower Eight     | -.463  |
| Lower Quartile  | -.174  |
| Median          | -.014  |
| Upper Quartile  | .211   |
| Upper Eight     | .436   |
| Upper Sixteenth | .706   |
| Maximum         | 1.246  |

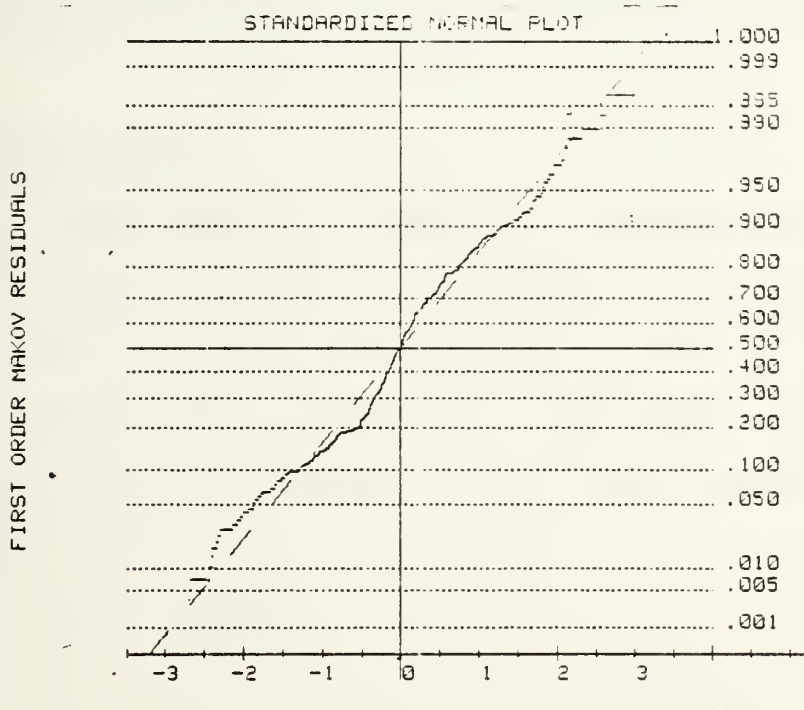


Figure 41. Standardized normal plot of first order Markov residuals from logged rainfall anomalies of data set RN



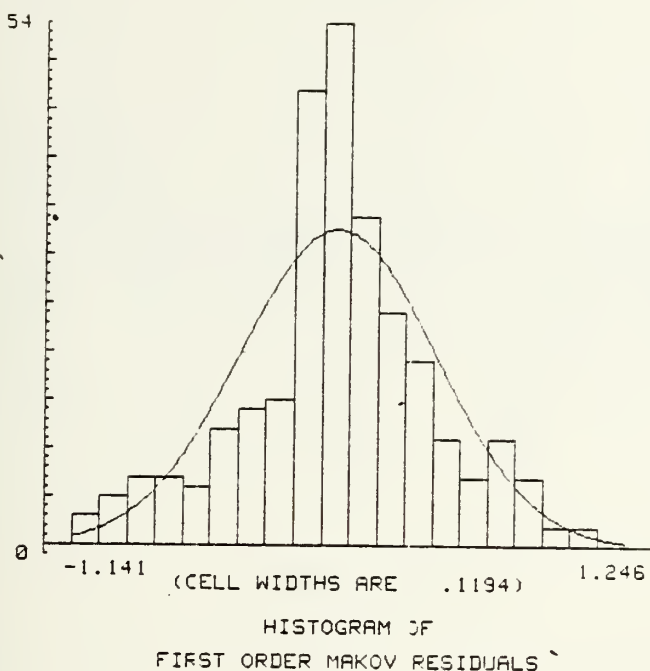


Figure 42. Histogram of first order Markov residuals from logged rainfall anomalies of data set RN

## 2. Winter Series

As stated above, the number of summer months with zeros indicated that a look at the winter months only might be worthwhile. The Figures 43 (Winter months), 44 (Correlogram), and 45 (Partial correlogram), which deal only with winter months, still indicate a first order autoregressive Markov model as;

$$\tilde{z}_t'' = .218\tilde{z}_{t-1}'' + a_{t-1} \quad \text{III.5}$$





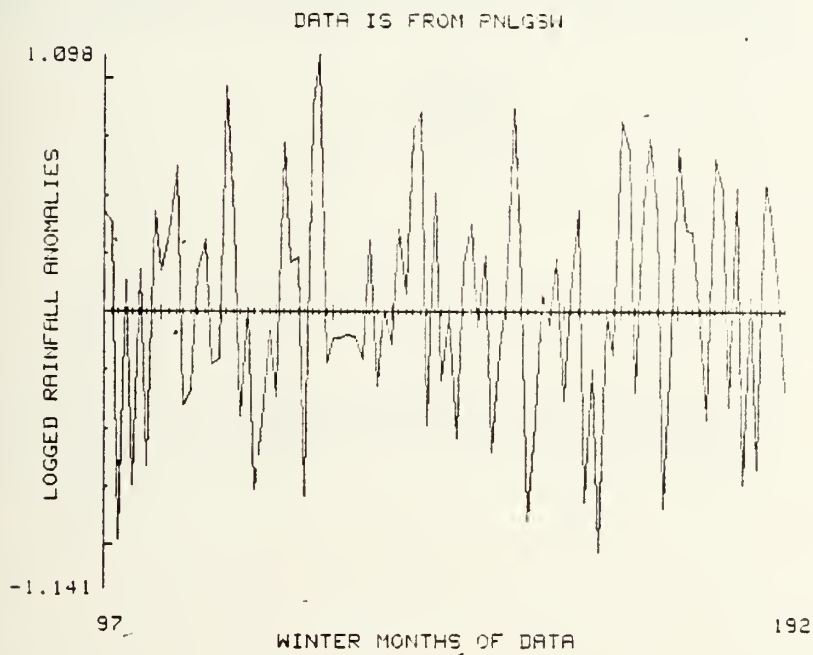
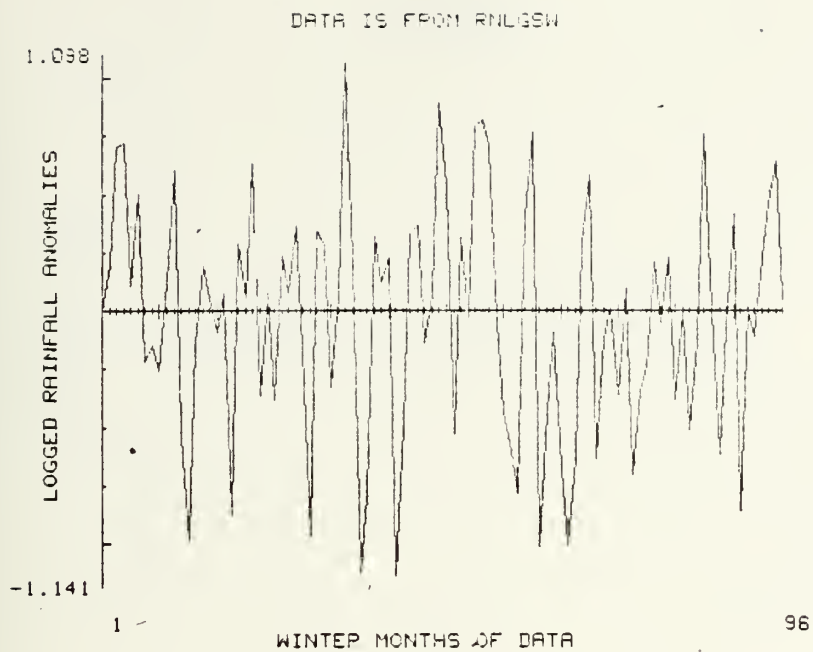


Figure 43. Winter months only of logged rainfall anomalies of data set RN



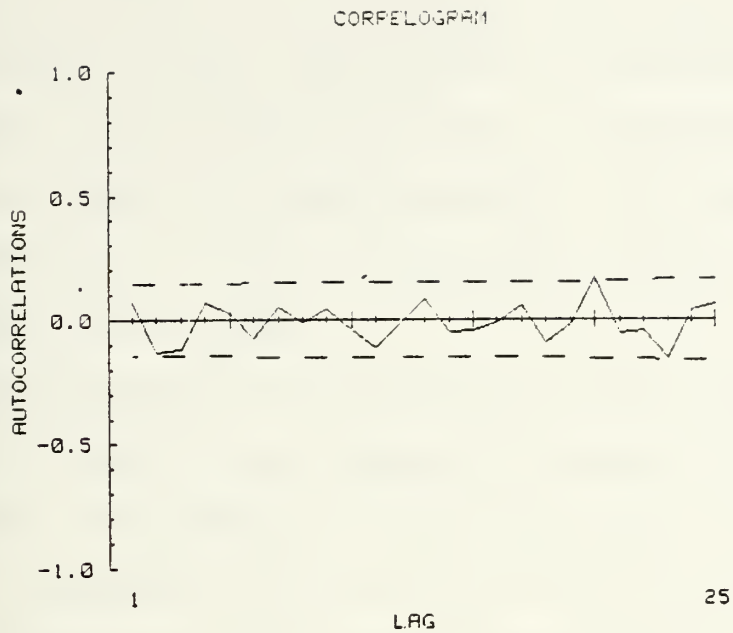


Figure 44. Correlogram of winter months only of logged rainfall anomalies from data set RN

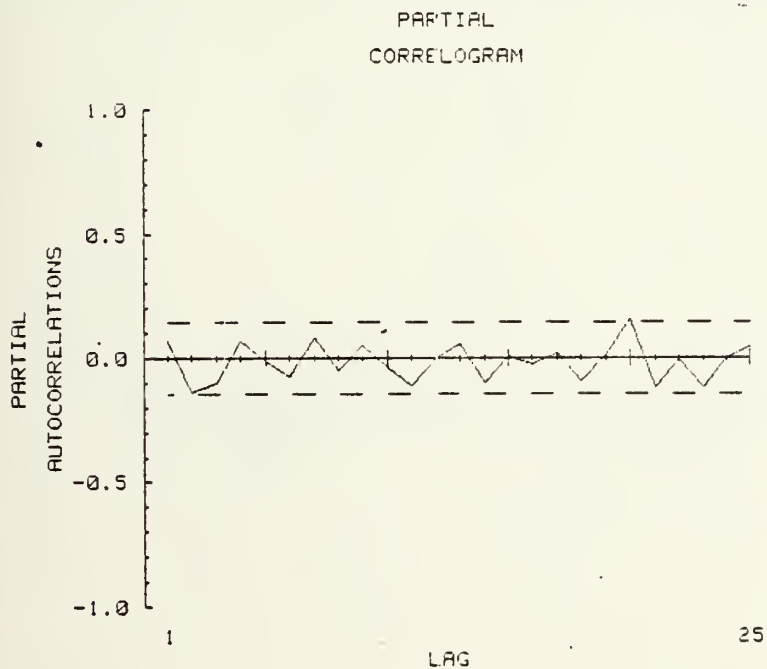


Figure 45. Partial autocorrelations of winter months only logged rainfall anomalies from data set RN



Now, as with the full twelve month model, a look at the residuals yields the Figures 46 (Residuals), 47 (Correlogram), and 48 (Lag one plot), 49 (Residuals vs. lag one), and Table 19 (Turning points). It appears that the residuals are, in fact, independent. This is similar to the twelve month model.

The question of the normality of the residuals is addressed by Table 19 and Figures 50 (Normal plot) and 51 (Histogram). The results of these plots and a basic chi-squared goodness of fit of 22.11 with 17 degrees of freedom indicate that this winter month data set is much more normal than was its twelve month counterpart. This chi-squared value is significant at the .181 level.

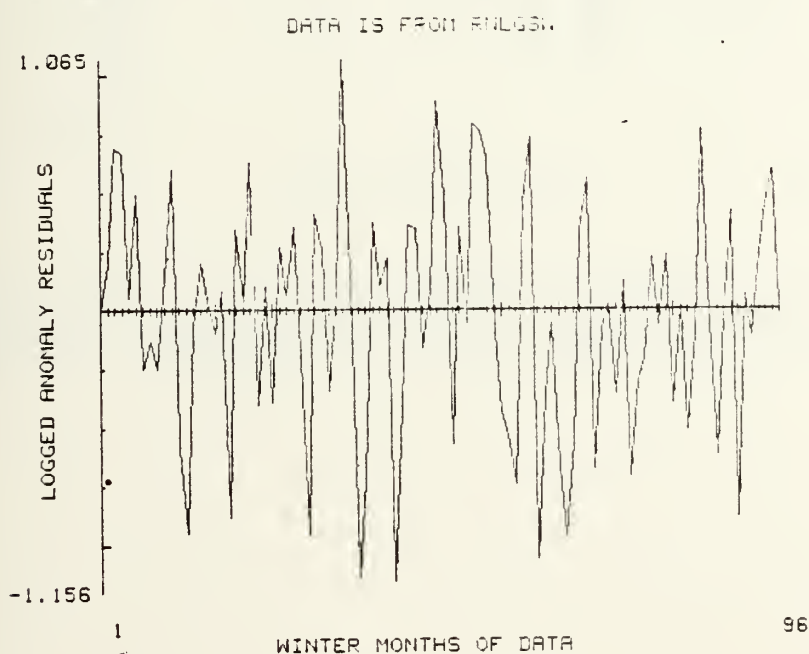


Figure 46a. First order Markov residuals of logged rainfall anomalies for winter months only of data set RN. Years 1 - 12



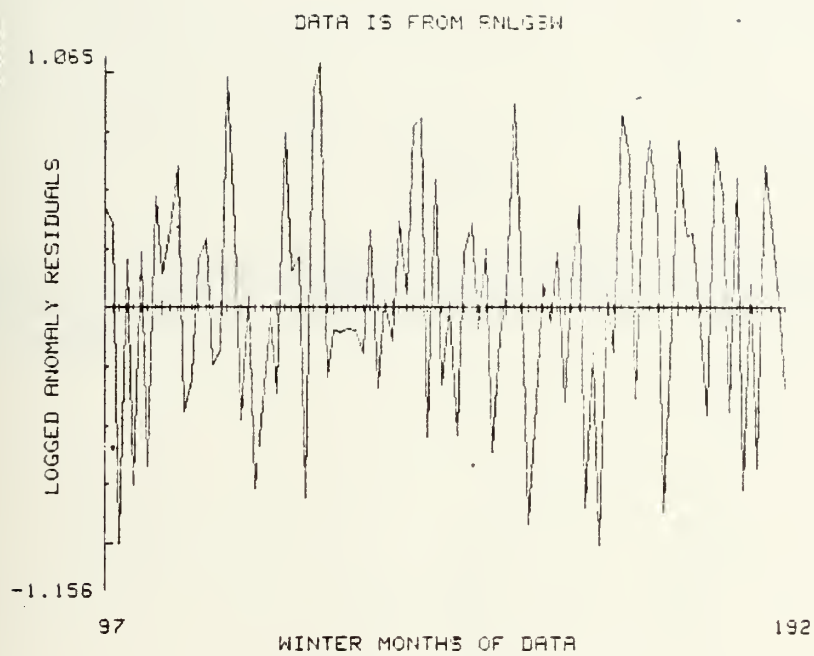


Figure 46b. First order Markov residuals of logged rainfall anomalies for winter months only of data set RN. Years 12 - 24





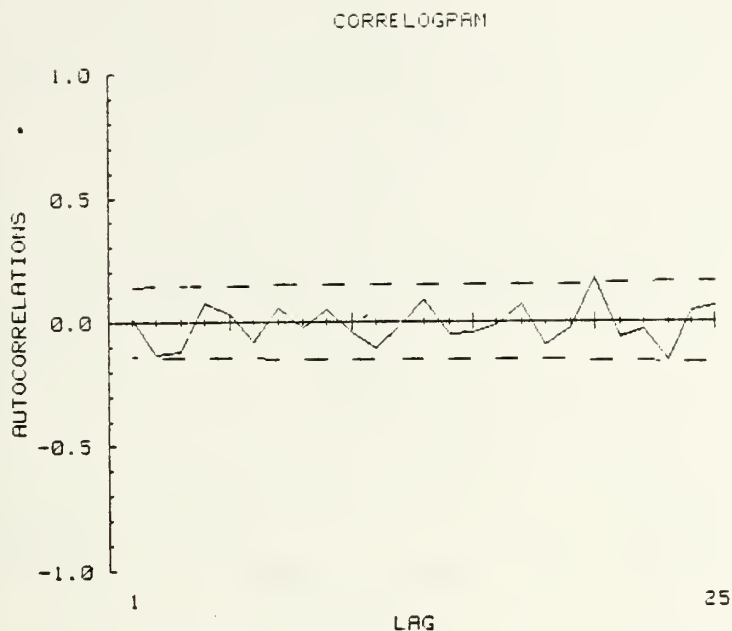


Figure 47. Correlogram of first order Markov residuals of logged rainfall anomalies for winter months only of data set RN

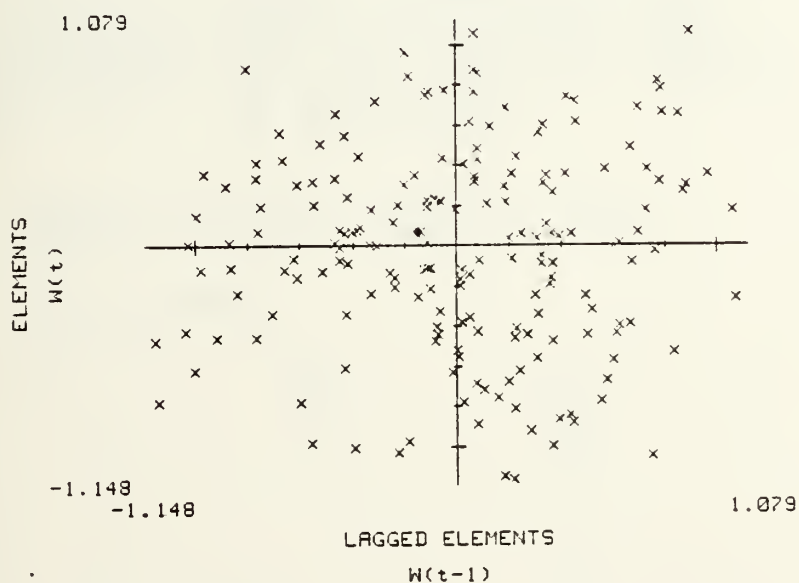


Figure 48. Lag one plot of first order Markov residuals from logged rainfall anomalies for winter months only of data set RN



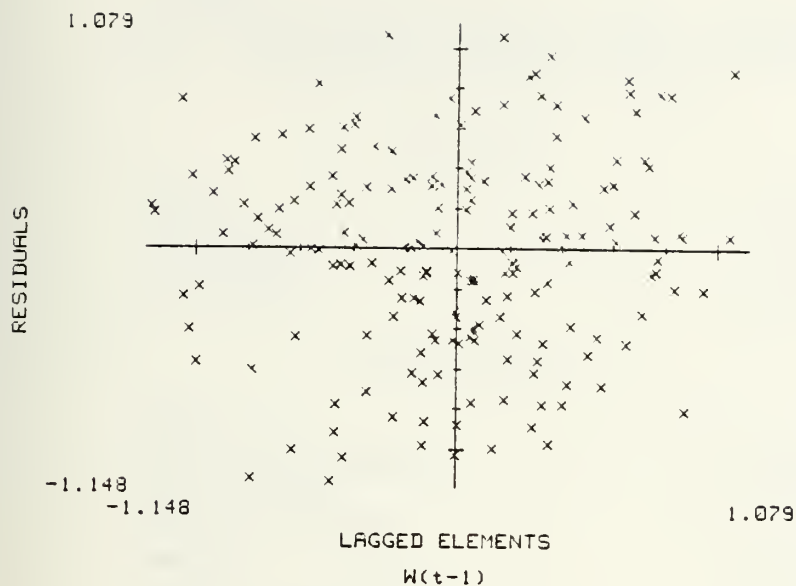


Figure 49. First order Markov residuals versus lag one data point from lagged rainfall anomalies of winter month only data from data set RN

TABLE 19

ACTUAL AND EXPECTED NUMBER OF TURNING  
POINTS AND ACTUAL AND EXPECTED PHASE  
FREQUENCIES FROM THE FIRST ORDER MARKOV  
RESIDUALS OF THE LOGGED RAINFALL ANOMALIES  
OF DATA SET RN

NUMBER OF TURNING POINTS = 129  
 $E[P] = 126.667$        $V[P] = 15.84$

PHASE LENGTHS

| D | OBS. | $E[*]$ |
|---|------|--------|
| 1 | 82   | 78.8   |
| 2 | 38   | 34.5   |
| 3 | 7    | 9.9    |
| 4 | 1    | 2.1    |
| 5 | 1    | .4     |
| 6 | 0    | 0.0    |
| 7 | 0    | 0.0    |



TABLE 20

GENERAL STATISTICS OF FIRST ORDER  
MARKOV RESIDUALS FROM LOGGED RAINFALL  
ANOMALIES OF DATA SET RN

## Moments

|          |       |
|----------|-------|
| Mean     | -.001 |
| Variance | .247  |
| Skewness | -.161 |
| Kurtosis | -.521 |

## Percentiles

|                 |        |
|-----------------|--------|
| Maximum         | -1.148 |
| Lower Sixteenth | -.845  |
| Lower Eight     | -.616  |
| Lower Quartile  | -.368  |
| Median          | .025   |
| Upper Quartile  | .339   |
| Upper Eight     | .591   |
| Upper Sixteenth | .762   |
| Maximum         | 1.078  |

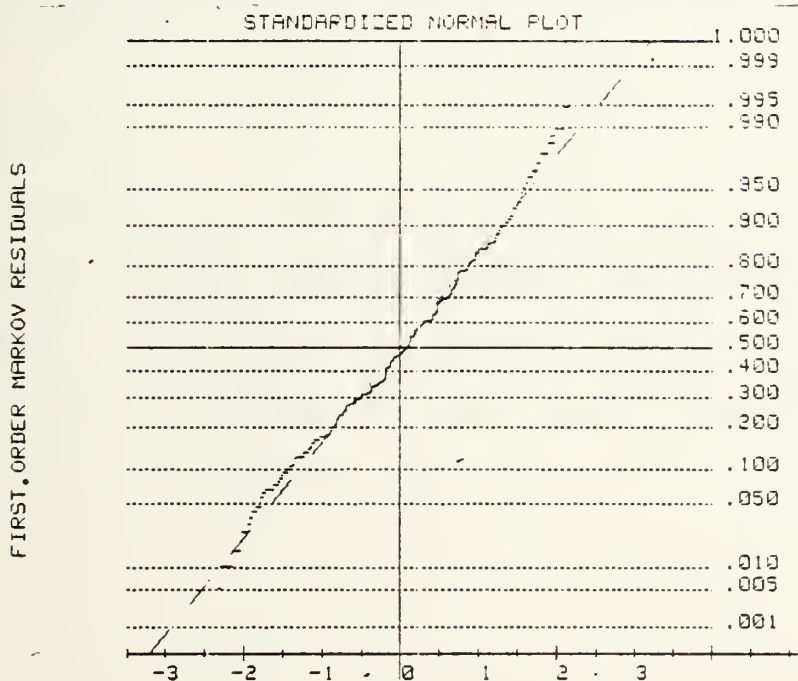


Figure 50. Standardized normal plot of first order Markov residuals from logged rainfall anomalies of winter months only from data set RN



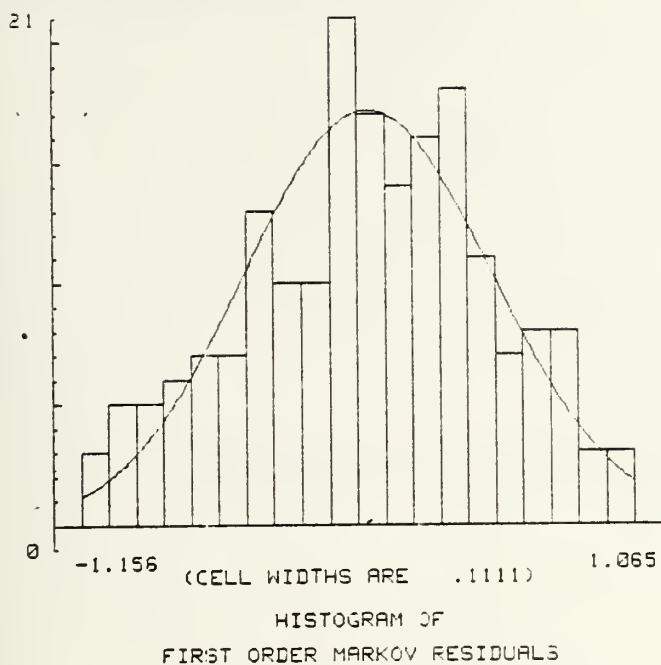


Figure 51. Histogram of first order Markov residuals from logged rainfall anomalies of winter month only of data set RN

#### C. DATA SET FL

As in the previous section on the data sets, the analysis of section III.B above carries forward fairly well to data sets FL and SC. This section, and the following, contain only the Figures and Tables corresponding to those in the previous section on data set RN.





# 1. Twelve Month Series

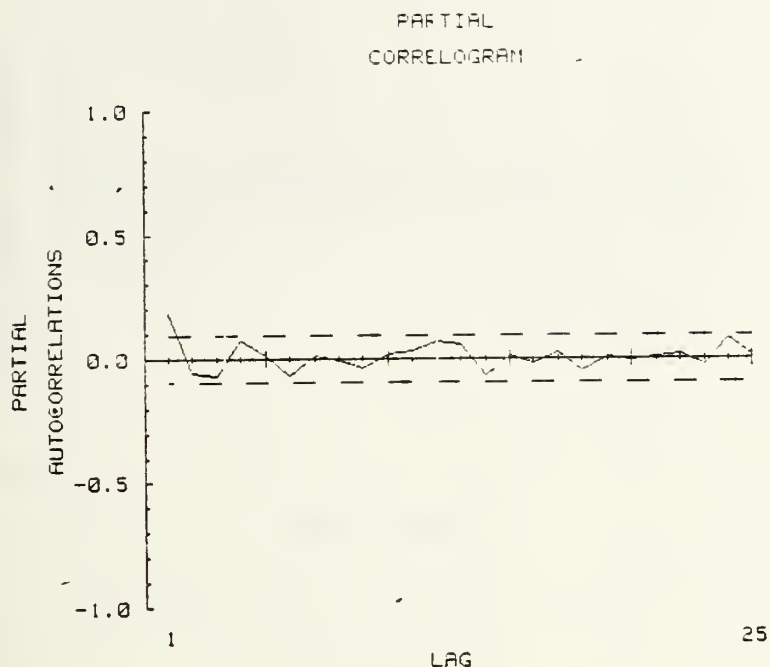


Figure 52. Partial correlogram of the logged rainfall anomalies for data set FL

TABLE 21

ESTIMATED PARTIAL AUTOCORRELATIONS FOR  
LOGGED RAINFALL ANOMALIES OF DATA SET FL

| LAG | VALUE | LAG | VALUE |
|-----|-------|-----|-------|
| 1   | .185  | 14  | -.072 |
| 2   | -.057 | 15  | .011  |
| 3   | -.069 | 16  | -.019 |
| 4   | .077  | 17  | .027  |
| 5   | .016  | 18  | -.048 |
| 6   | -.068 | 19  | .006  |
| 7   | .010  | 20  | -.007 |
| 8   | -.008 | 21  | .003  |
| 9   | -.040 | 22  | .018  |
| 10  | .022  | 23  | -.025 |
| 11  | .031  | 24  | .081  |
| 12  | .067  | 25  | .012  |
| 13  | .055  |     |       |

$$\tilde{z}'_t = .185\tilde{z}'_{t-1} + a_t$$

III.6



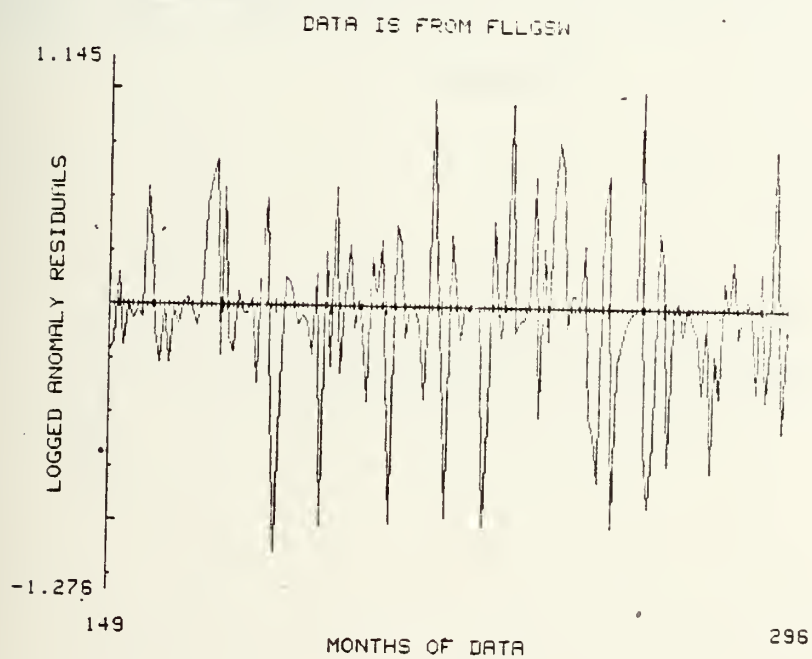
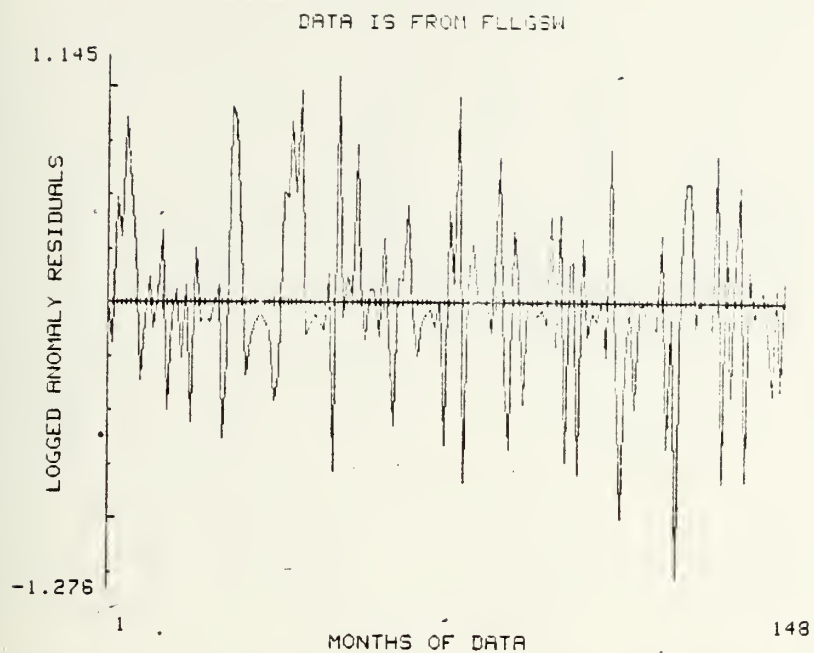


Figure 53a. First order Markov residuals from logged rainfall anomalies of data set FL. Months 1 - 296



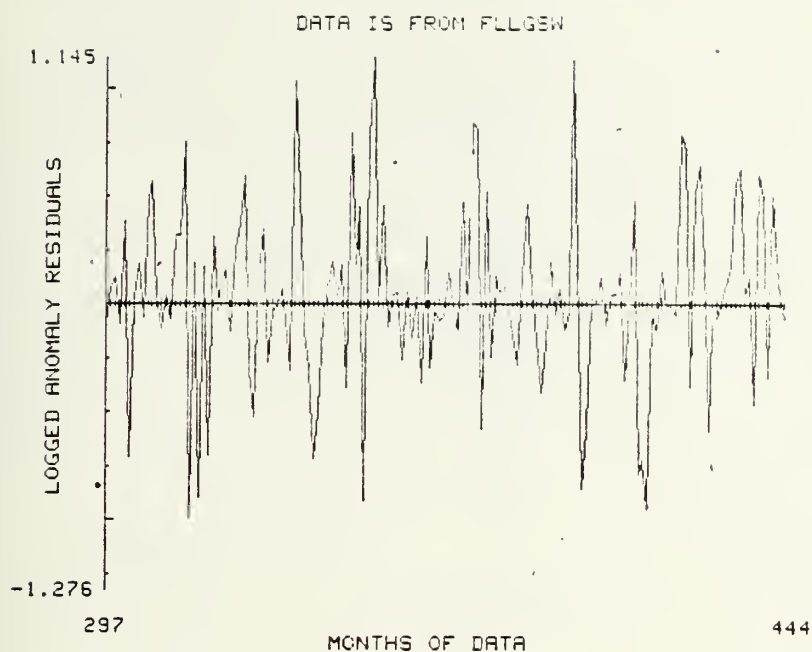


Figure 53b. First order Markov residuals from logged rainfall anomalies of data set FL. Months 297 - 444.

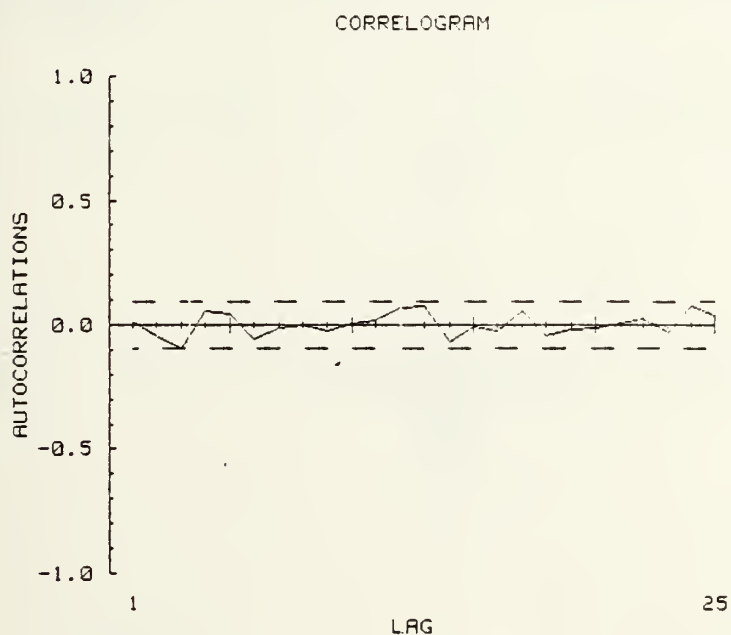


Figure 54. Autocorrelations of residuals from first order Markov process applied to the logged rainfall anomalies of data set FL



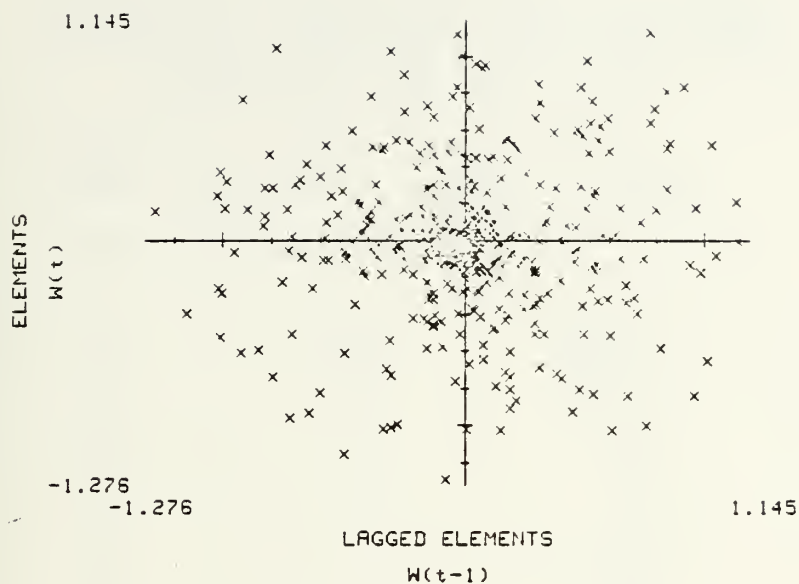


Figure 55. Lag one plot of first order Markov residuals from logged rainfall anomalies of data set FL

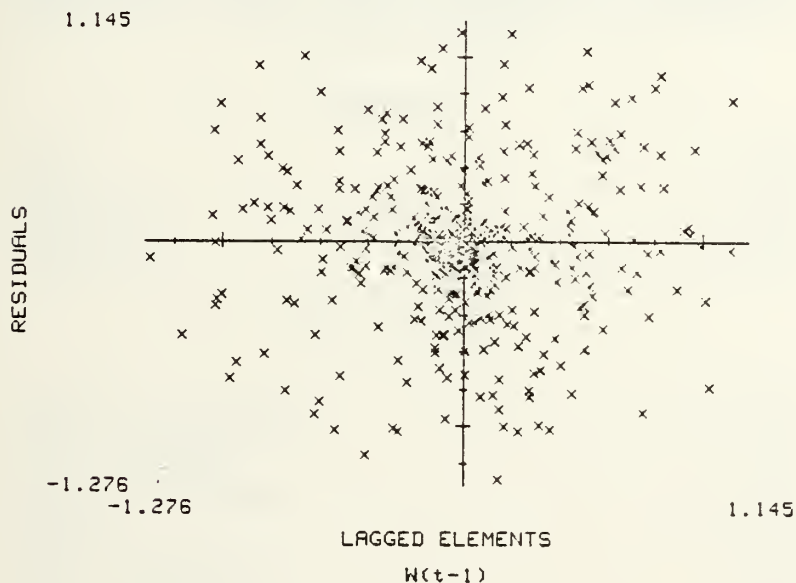


Figure 56. First order Markov residuals versus lag one data points from logged rainfall anomalies of data set FL





TABLE 22

ACTUAL AND EXPECTED NUMBER OF TURNING  
POINTS AND ACTUAL AND EXPECTED PHASE  
FREQUENCIES FROM THE FIRST ORDER MARKOV  
RESIDUALS FROM DATA SET FL

NUMBER OF TURNING POINTS = 294  
E[P] = 294.667      V[P] = 15.93

## PHASE LENGTHS

| D      | OBS. | E[*]  |
|--------|------|-------|
| 1      | 188  | 183.8 |
| 2      | 79   | 80.7  |
| 3      | 19   | 23.2  |
| 4      | 6    | 5.0   |
| 5      | 0    | .9    |
| 6      | 1    | .1    |
| 7      | 1    | 0.0   |
| 8      | 0    | 0.0   |
| 9      | 0    | 0.0   |
| 10     | 0    | 0.0   |
| TOTALS | 294  | 293.7 |

TABLE 23

GENERAL STATISTICS OF FIRST ORDER MARKOV  
RESIDUALS FROM LOGGED RAINFALL ANOMALIES  
OF DATA SET FL

## Moments

|          |       |
|----------|-------|
| Mean     | .000  |
| Variance | .163  |
| Skewness | -.044 |
| Kurtosis | .717  |

## Percentiles

|                 |        |
|-----------------|--------|
| Minimum         | -1.276 |
| Lower Sixteenth | -.702  |
| Lower Eight     | -.426  |
| Lower Quartile  | -.156  |
| Median          | -.012  |
| Upper Quartile  | .184   |
| Upper Eight     | .481   |
| Upper Sixteenth | .648   |
| Maximum         | 1.124  |



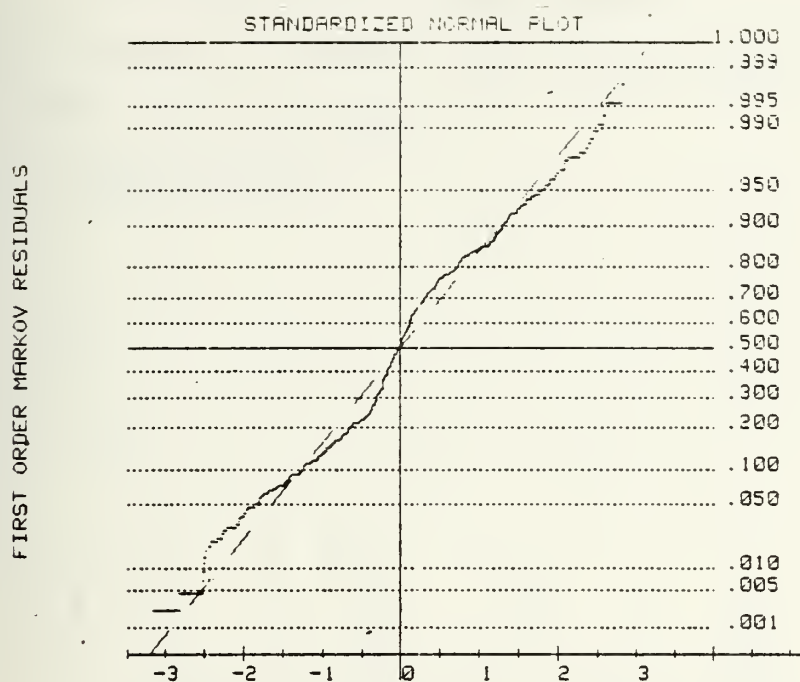


Figure 57. Standardized normal plot of first order Markov residuals from logged rainfall anomalies of data set FL

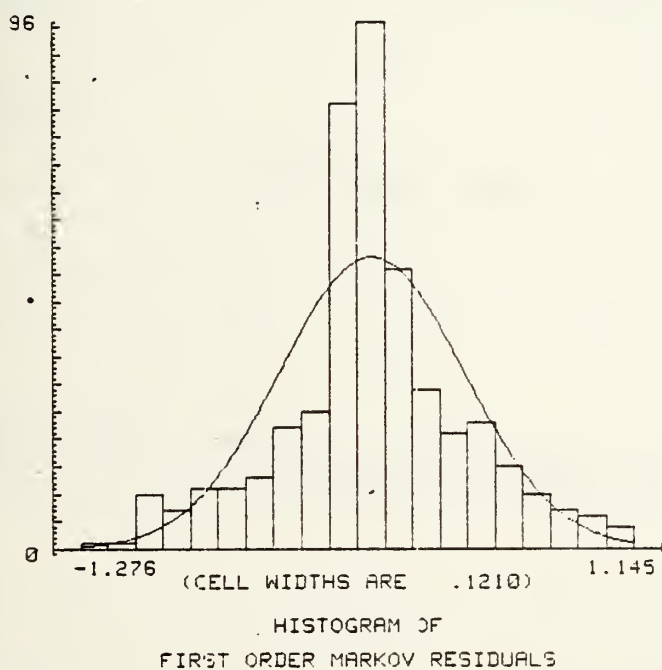


Figure 58. Histogram of first order Markov residuals from logged rainfall anomalies of data set FL



This data set yielded a chi-square value of 107.66 for 17 degrees of freedom. The significance of the value is in the zero plus range.

## 2. Winter Series

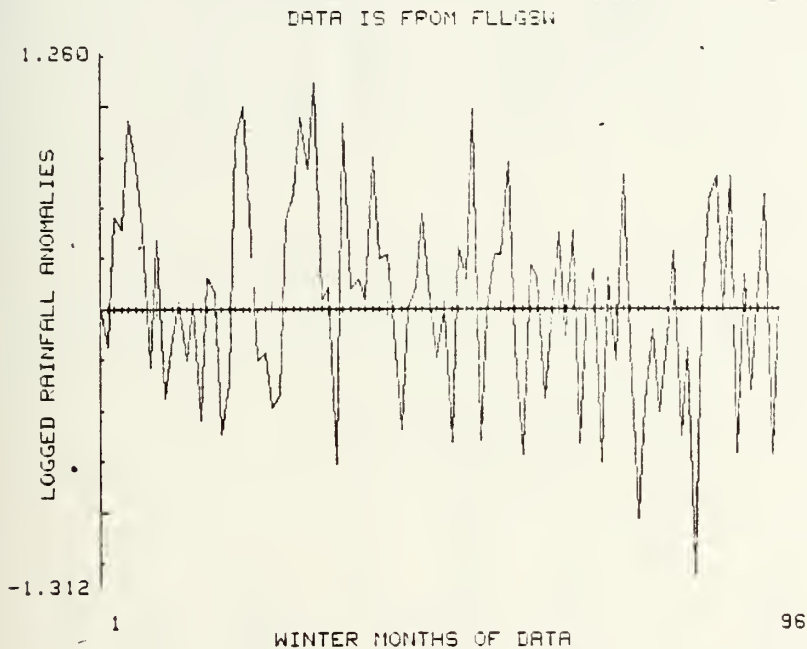
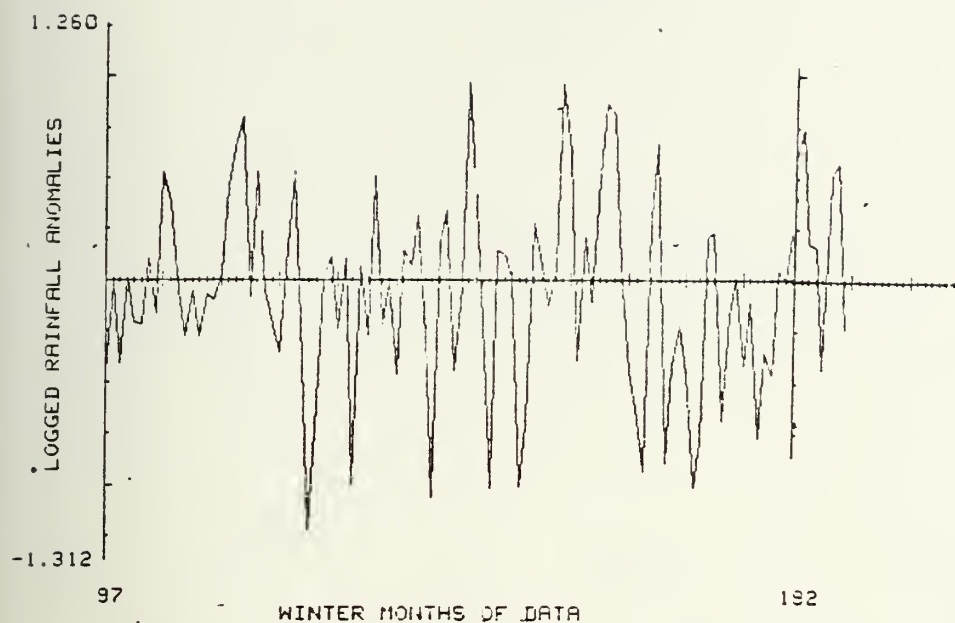


Figure 59a. Years 1 - 12 of winter months only of logged rainfall anomalies of data set FL



DATA IS FROM FLLGSW



DATA IS FROM FLLGSW

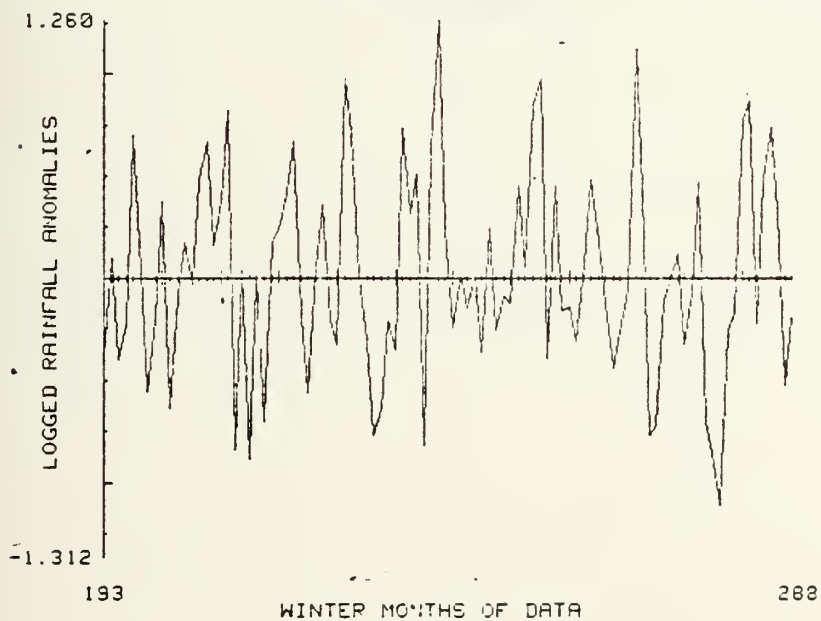


Figure 59b. Years 13 - 37 of winter months only, logged rainfall anomalies of data set FL





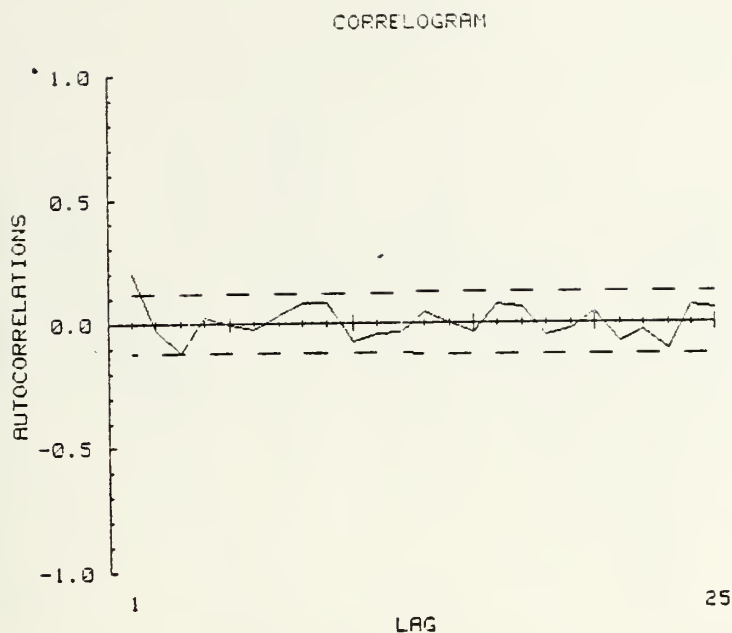


Figure 60. Correlogram of winter months only, logged rainfall anomalies from data set FL

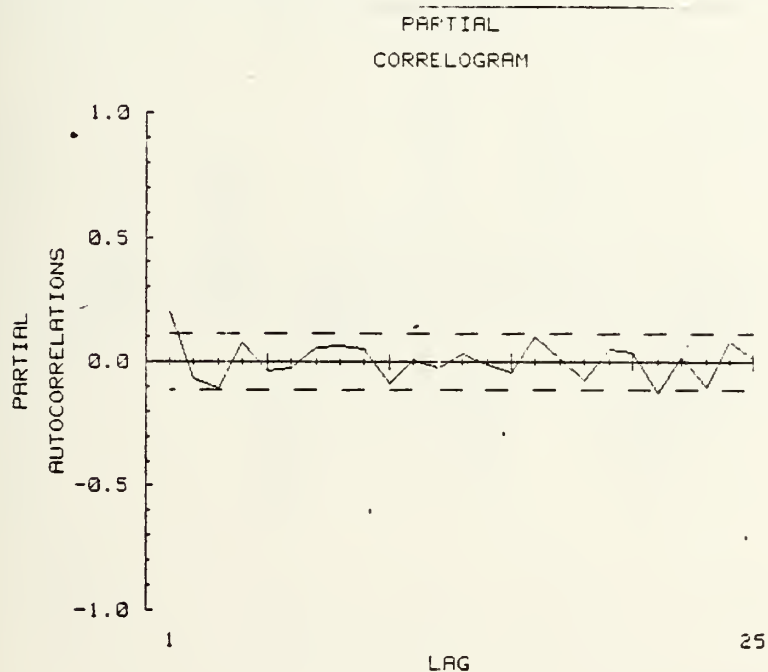


Figure 61. Partial correlogram of winter months only, logged rainfall anomalies from data set FL

These displays indicate a model like

$$\tilde{z}_t'' = .199\tilde{z}_{t-1}'' + a_t.$$

III.7



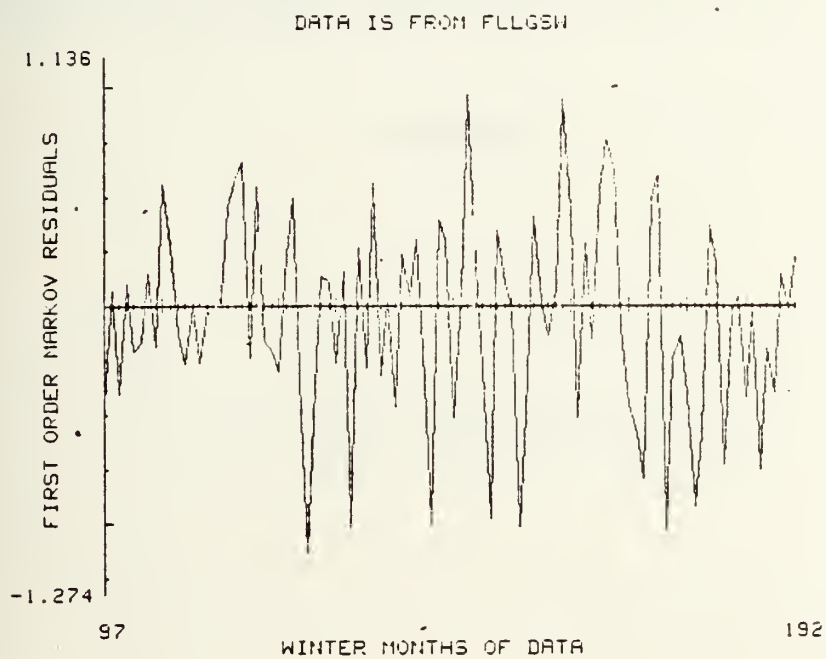
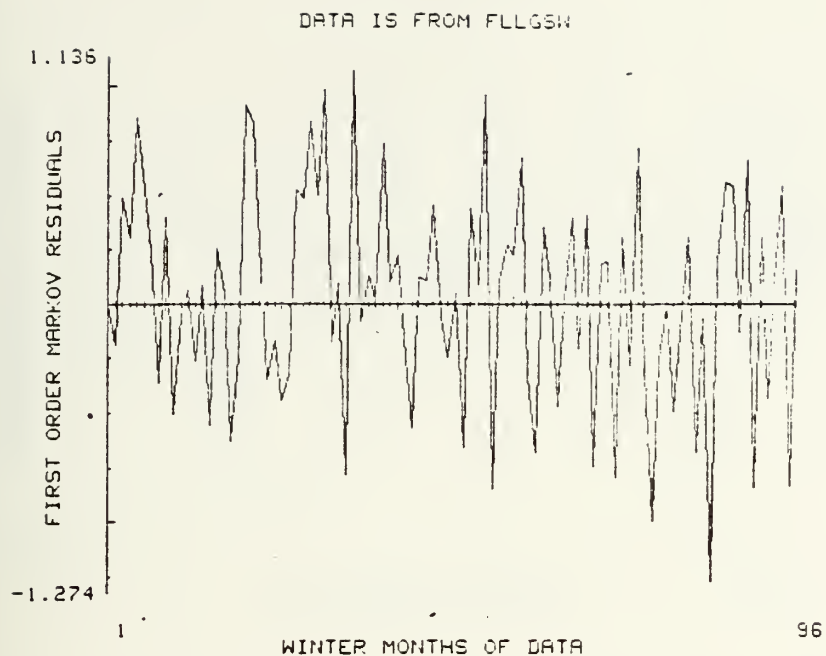


Figure 62a. Years 1 - 24, first order Markov residuals of logged rainfall anomalies for winter months only, data set FL



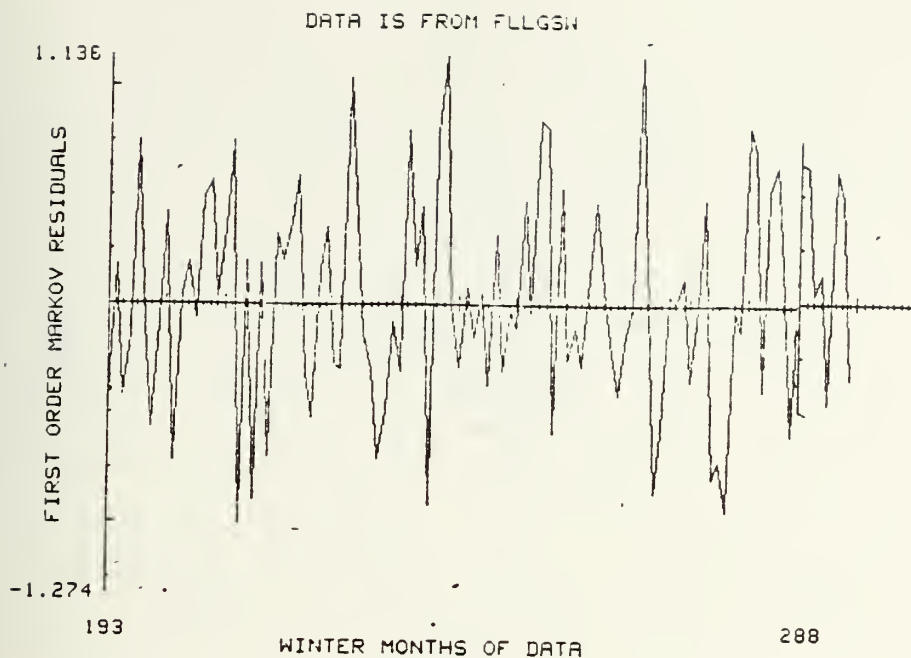


Figure 62b. Years 25 - 37, first order Markov residuals of logged rainfall anomalies for winter months only, data set FL

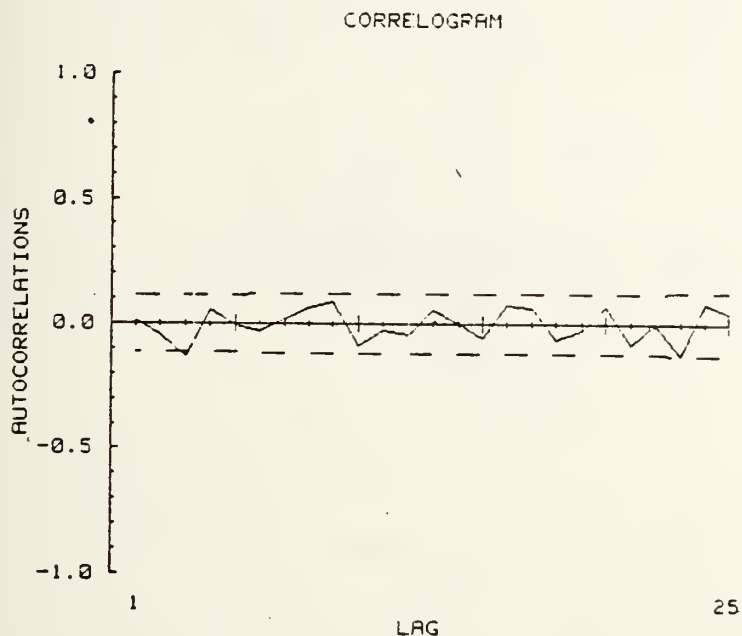


Figure 63. Correlogram of first order Markov residuals of lagged rainfall anomalies from winter months only, data set FL



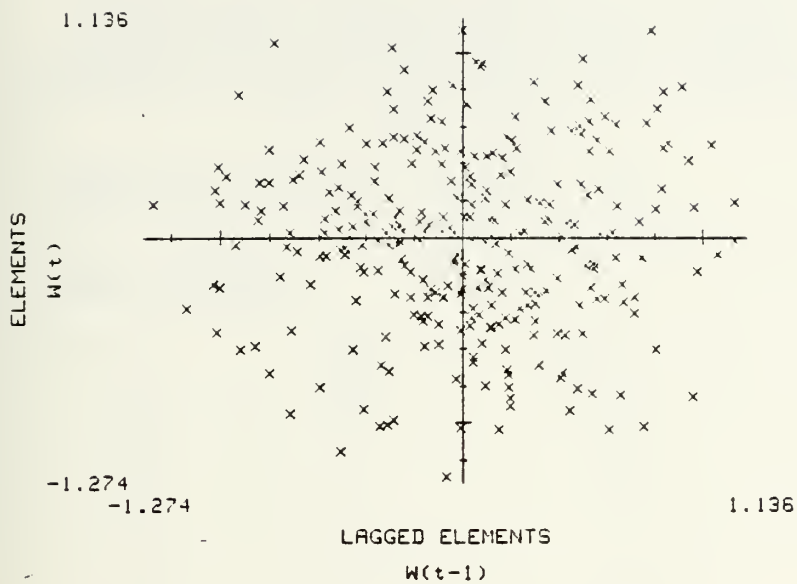


Figure 64. Lag one plot of first order Markov residuals from logged rainfall anomalies of winter months only, data set FL

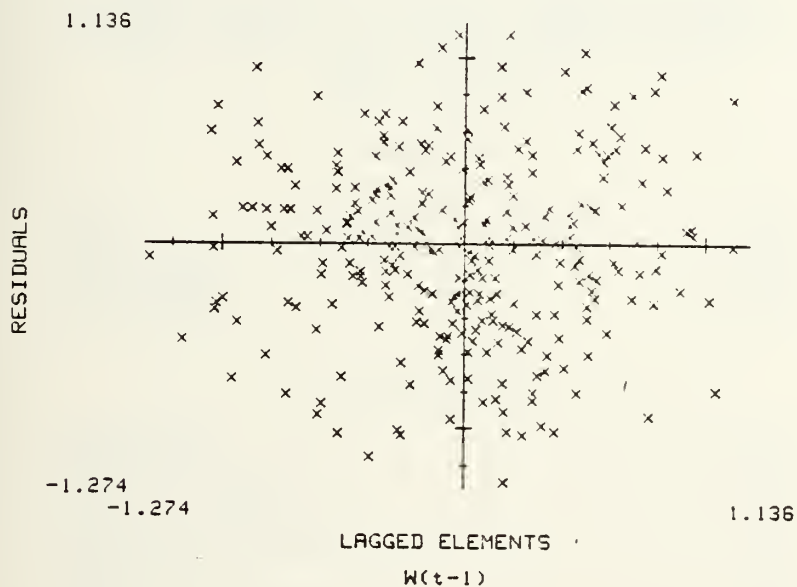


Figure 65. First order Markov residuals versus lag one data point from logged rainfall anomalies of winter months only, data set FL





TABLE 24

ACTUAL AND EXPECTED NUMBER OF TURNING  
POINTS AND ACTUAL AND EXPECTED PHASE  
FREQUENCIES FROM THE FIRST ORDER MARKOV  
RESIDUALS OF THE LOGGED RAINFALL ANOMALIES  
OF DATA SET FL

NUMBER OF TURNING POINTS = 209  
E[P] = 196                      V[P] = 15.902

## PHASE LENGTHS

| D | OBS. | E[*]  |
|---|------|-------|
| 1 | 138  | 122.1 |
| 2 | 59   | 53.5  |
| 3 | 8    | 15.4  |
| 4 | 2    | 3.3   |
| 5 | 1    | .6    |
| 6 | 0    | .1    |
| 7 | 0    | 0.0   |
| 8 | 0    | 0.0   |

TABLE 25

GENERAL STATISTICS OF FIRST ORDER  
MARKOV RESIDUALS FROM LOGGED RAINFALL  
ANOMALIES OF WINTER MONTHS ONLY,  
DATA SET FL

## Moments

|          |       |
|----------|-------|
| Mean     | .000  |
| Variance | .234  |
| Skewness | -.079 |
| Kurtosis | -.376 |

## Percentiles

|                 |        |
|-----------------|--------|
| Minimum         | -1.274 |
| Lower Sixteenth | -.798  |
| Lower Eight     | -.570  |
| Lower Quartile  | -.315  |
| Median          | .011   |
| Upper Quartile  | .323   |
| Upper Eight     | .551   |
| Upper Sixteenth | .748   |
| Maximum         | 1.136  |



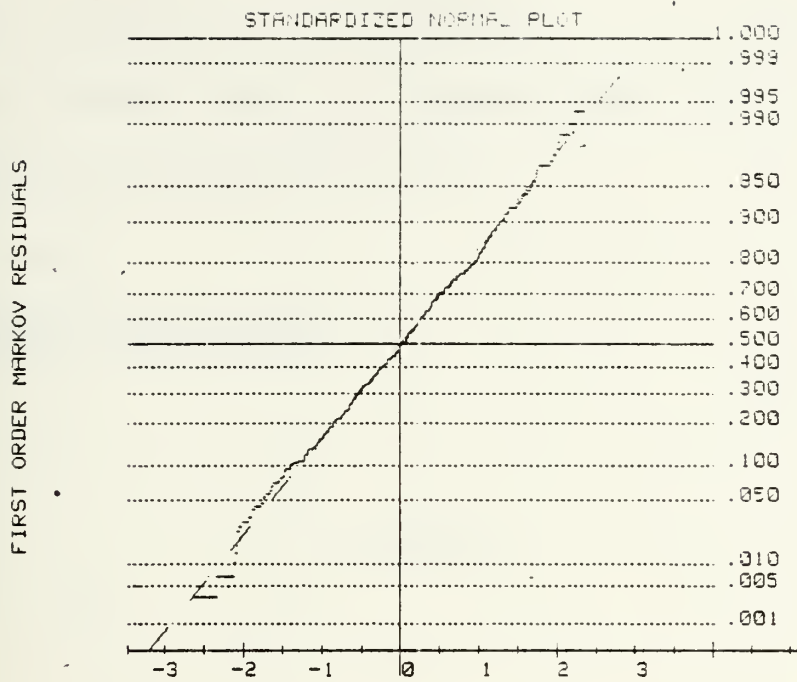


Figure 66. Standardized normal plot of first order Markov residuals from logged rainfall anomalies of winter months only, data set FL

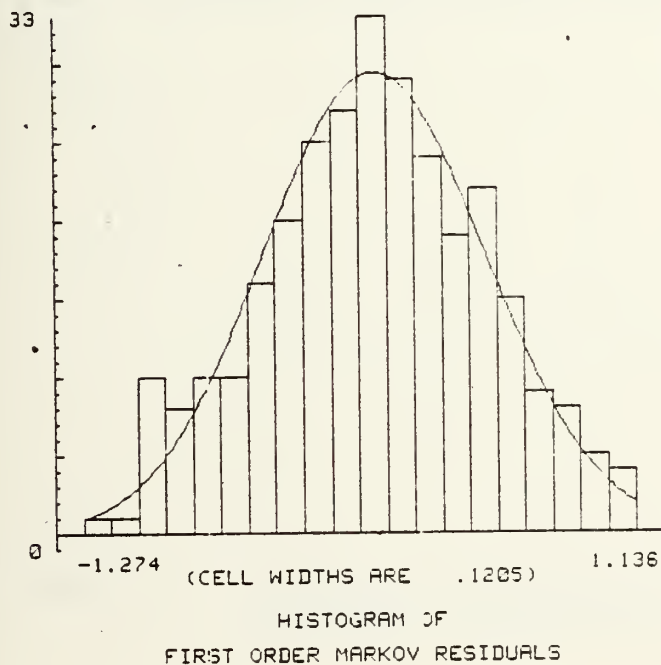


Figure 67. Histogram of first order Markov residuals from logged rainfall anomalies of winter months only, data set FL



The chi-squared was calculated at 15.35 for 17 degrees of freedom. This is a significance level of .570, thus indicating possible normality.

#### D. DATA SET SC

##### 1. Twelve Month Series

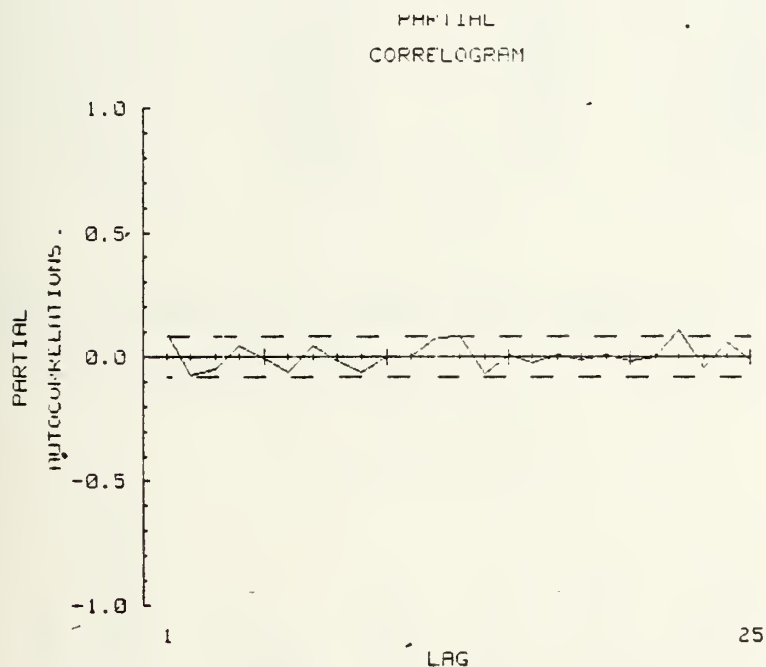


Figure 68. Partial correlogram of the logged anomalies for data set SC



TABLE 26

ESTIMATED PARTIAL AUTOCORRELATIONS  
FOR LOGGED RAINFALL ANOMALIES OF  
DATA SET SC

| LAG | VALUE | LAG | VALUE |
|-----|-------|-----|-------|
| 1   | .096  | 14  | -.066 |
| 2   | -.075 | 15  | .004  |
| 3   | -.053 | 16  | -.022 |
| 4   | .046  | 17  | .007  |
| 5   | -.004 | 18  | -.014 |
| 6   | -.061 | 19  | .004  |
| 7   | .042  | 20  | -.019 |
| 8   | -.016 | 21  | .002  |
| 9   | -.061 | 22  | .106  |
| 10  | -.000 | 23  | -.046 |
| 11  | -.001 | 24  | .059  |
| 12  | .079  | 25  | -.015 |
| 13  | .084  |     |       |

This information yields the model as

$$\tilde{z}'_t = .096\tilde{z}'_{t-1} + a_t.$$

III.12

DATA IS FROM SOLOSW

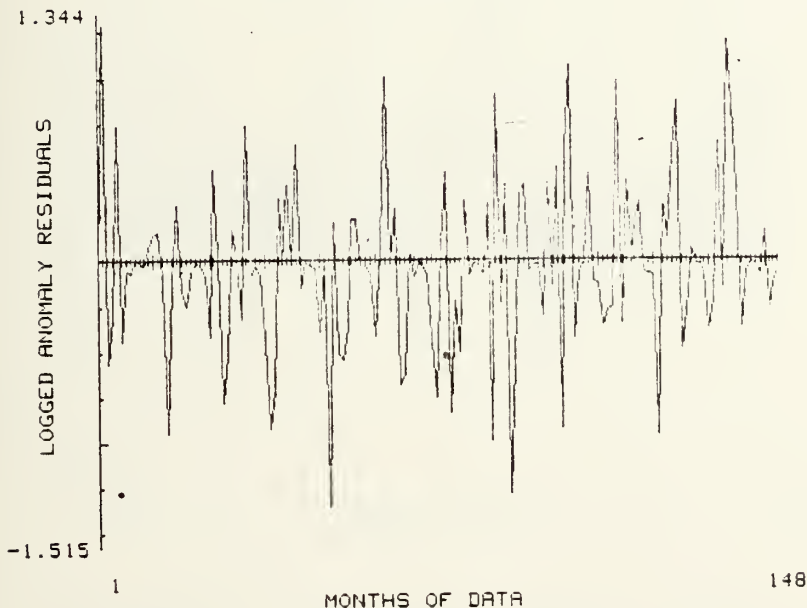


Figure 69a. First order Markov residuals from logged rainfall anomalies of data set SC. Months 1 - 148.





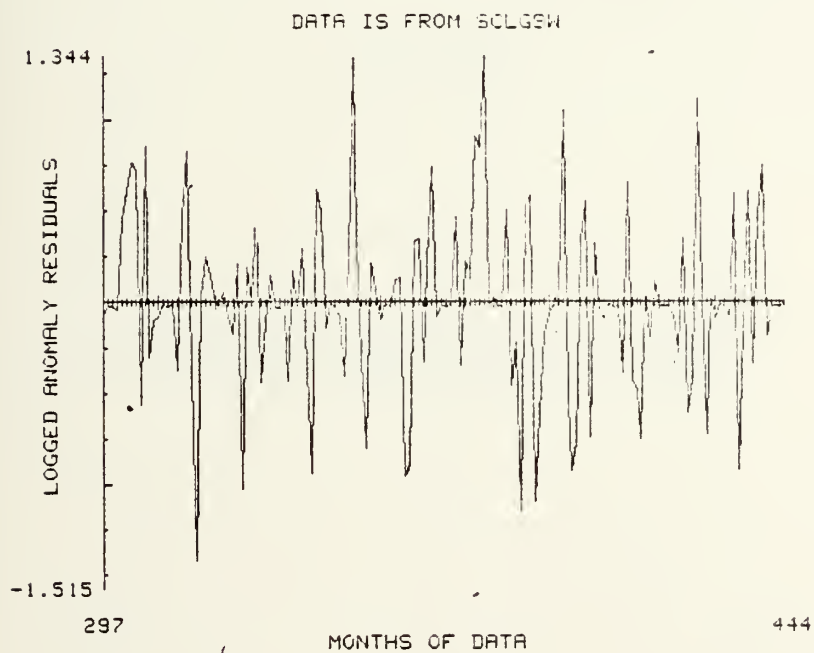
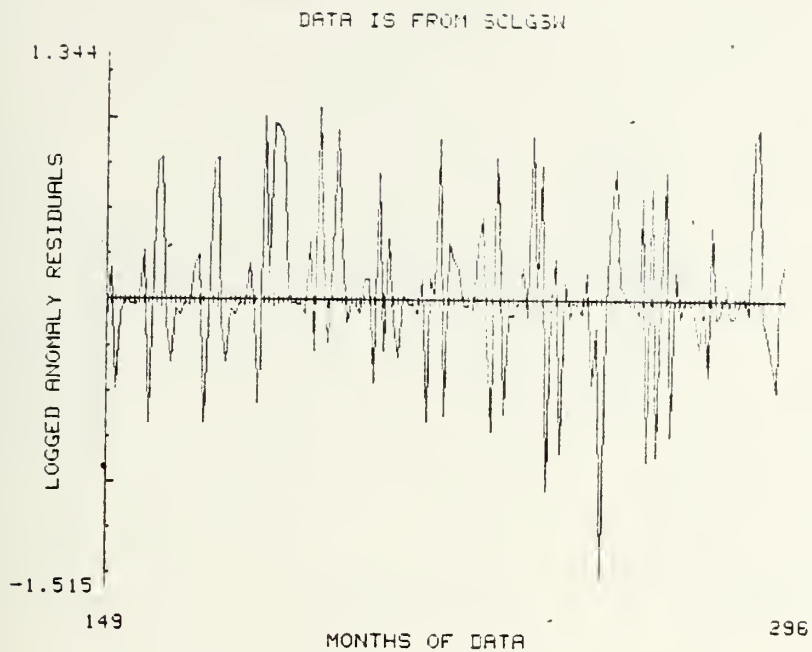


Figure 69b. First order Markov residual from logged rainfall anomalies of data set SC. Months 149 - 444



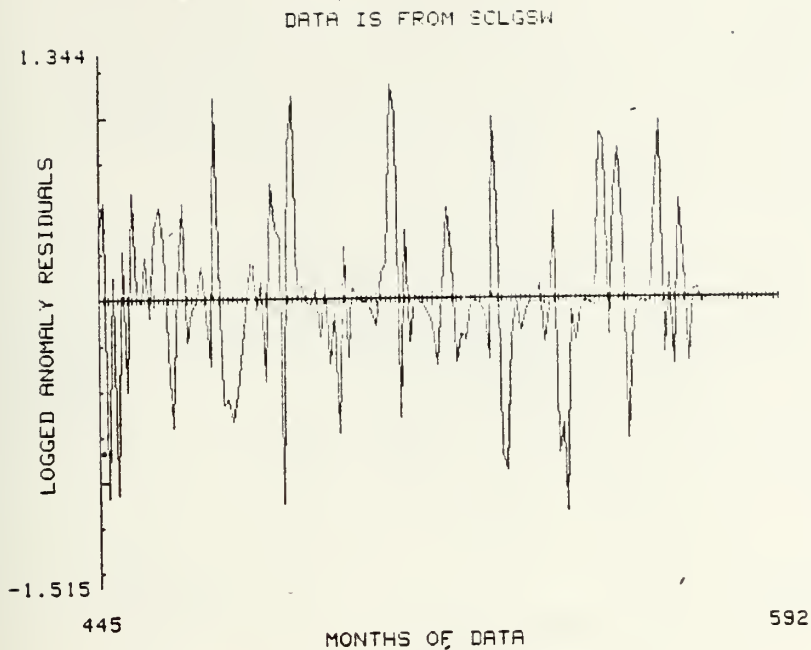


Figure 69c. First order Markov residual from logged rainfall anomalies of data set SC. Months 445 - 596.

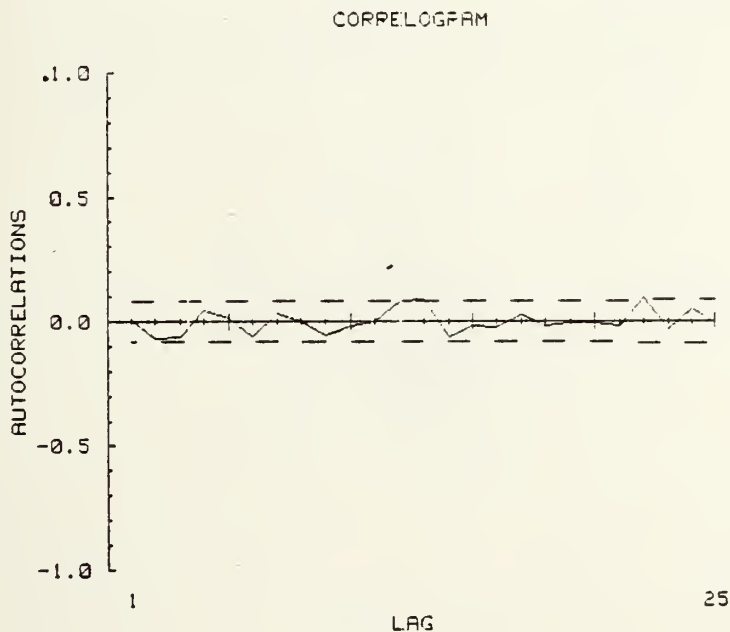


Figure 70. Autocorrelations of residuals from first order Markov process applied to the logged rainfall anomalies of data set SC



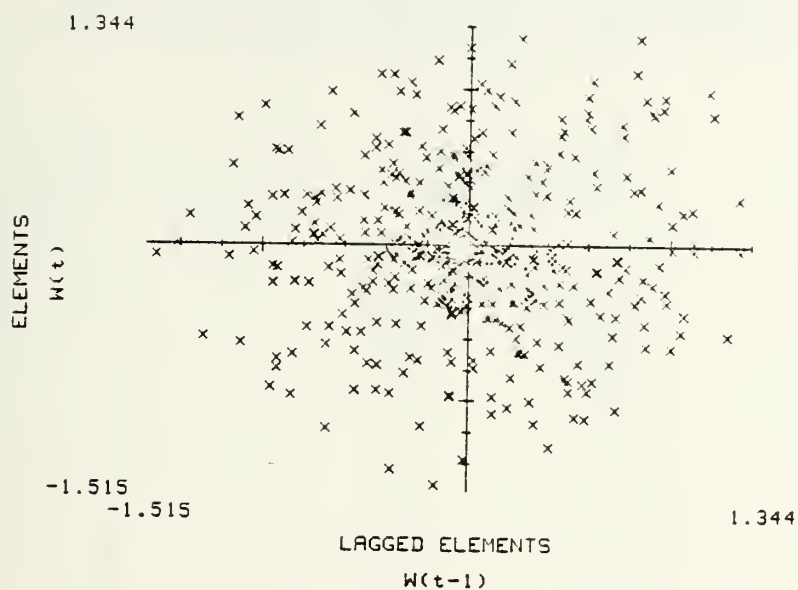


Figure 71. Lag one plot of first order Markov residuals from logged rainfall anomalies of data set SC

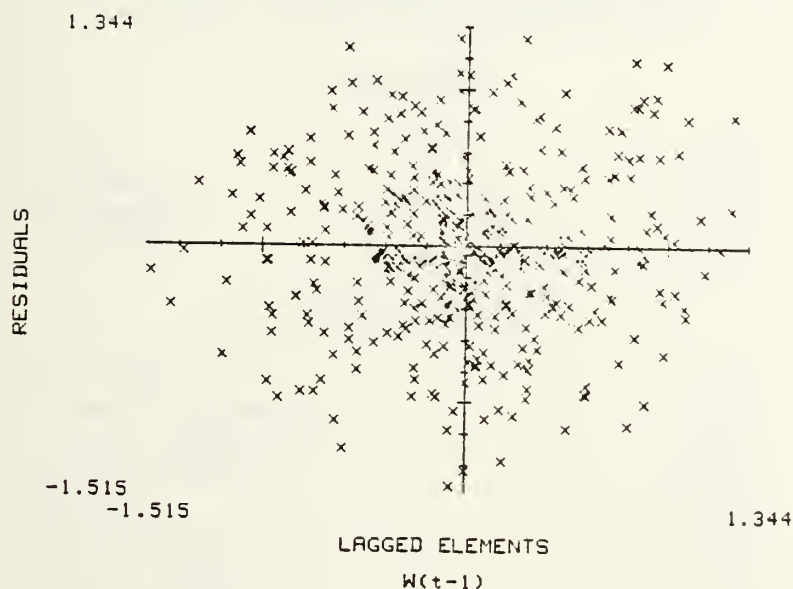


Figure 72. First order Markov residuals versus lag one data points from logged rainfall anomalies of data set SC



TABLE 27

ACTUAL AND EXPECTED NUMBER OF TURNING  
POINTS AND ACTUAL AND EXPECTED PHASE  
FREQUENCIES FROM THE FIRST ORDER MARKOV  
RESIDUALS OF DATA SET SC

NUMBER OF TURNING POINTS = 367  
E[P] = 382.667      V[P] = 15.9

## PHASE LENGTHS

| D      | OBS. | E[*]  |
|--------|------|-------|
| 1      | 226  | 238.8 |
| 2      | 95   | 104.9 |
| 3      | 34   | 30.1  |
| 4      | 8    | 6.6   |
| 5      | 2    | 1.2   |
| 6      | 1    | .2    |
| 7      | 1    | 0.0   |
| 8      | 0    | 0.0   |
| 9      | 0    | 0.0   |
| 10     | 0    | 0.0   |
| TOTALS | 367  | 381.7 |

TABLE 28

GENERAL STATISTICS OF FIRST ORDER  
MARKOV RESIDUALS FROM LOGGED  
RAINFALL ANOMALIES OF DATA SET SC

## Moments

|          |      |
|----------|------|
| Mean     | .000 |
| Variance | .209 |
| Skewness | .057 |
| Kurtosis | .805 |

## Percentiles

|                 |        |
|-----------------|--------|
| Minimum         | -1.515 |
| Lower Sixteenth | - .744 |
| Lower Eight     | - .462 |
| Lower Quartile  | - .205 |
| Median          | - .030 |
| Upper Quartile  | .207   |
| Upper Eight     | .506   |
| Upper Sixteenth | .783   |
| Maximum         | 1.344  |





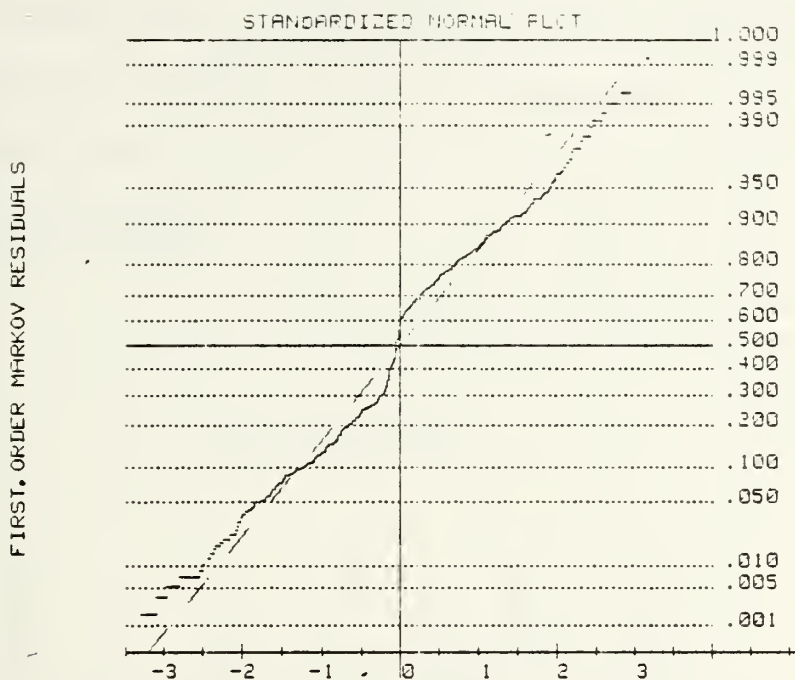


Figure 73. Standardized normal plot of first order Markov residuals from logged rainfall anomalies of data set SC

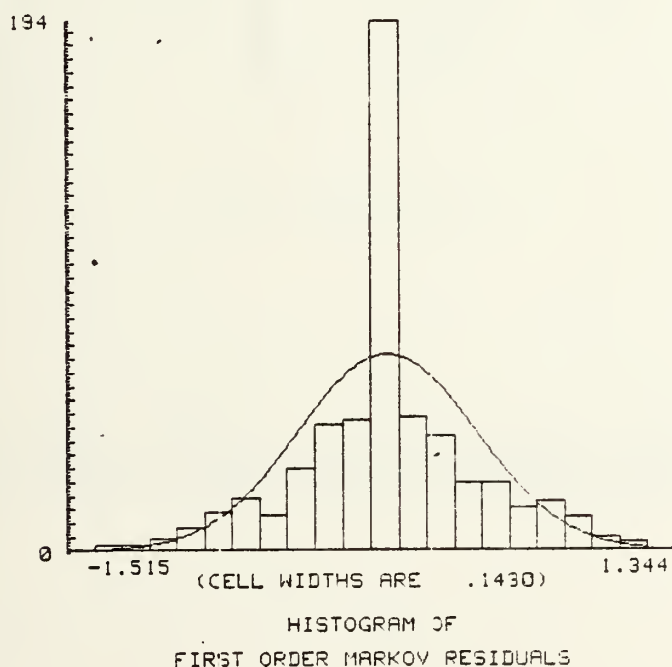


Figure 74. Histogram of first order Markov residuals from logged rainfall anomalies of data set SC



This data set yielded a chi-square value of 273.95 for 17 degrees of freedom. This is equivalent to a significance of zero plus.

## 2. Winter Series

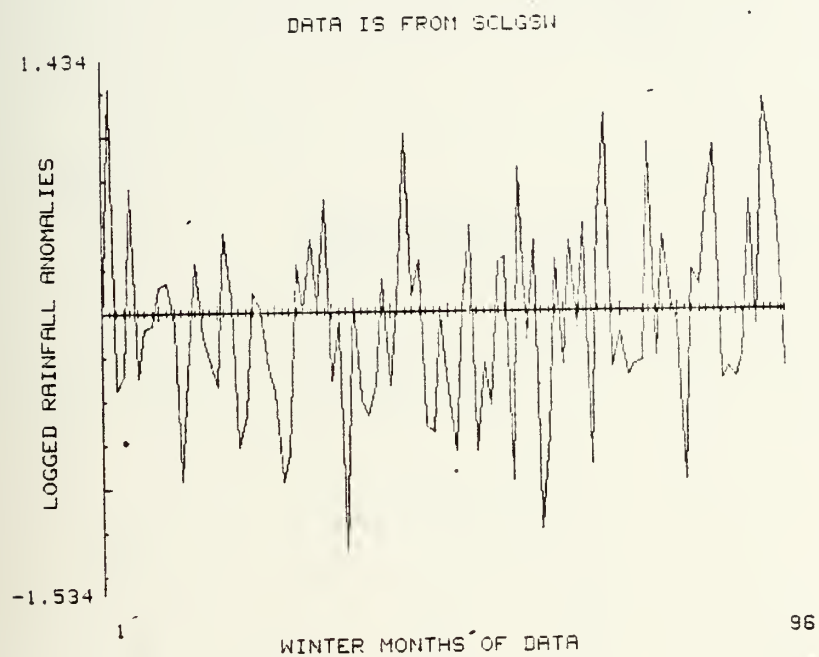


Figure 75a. Years 1 - 12 of winter months only, logged rainfall anomalies of data set SC.



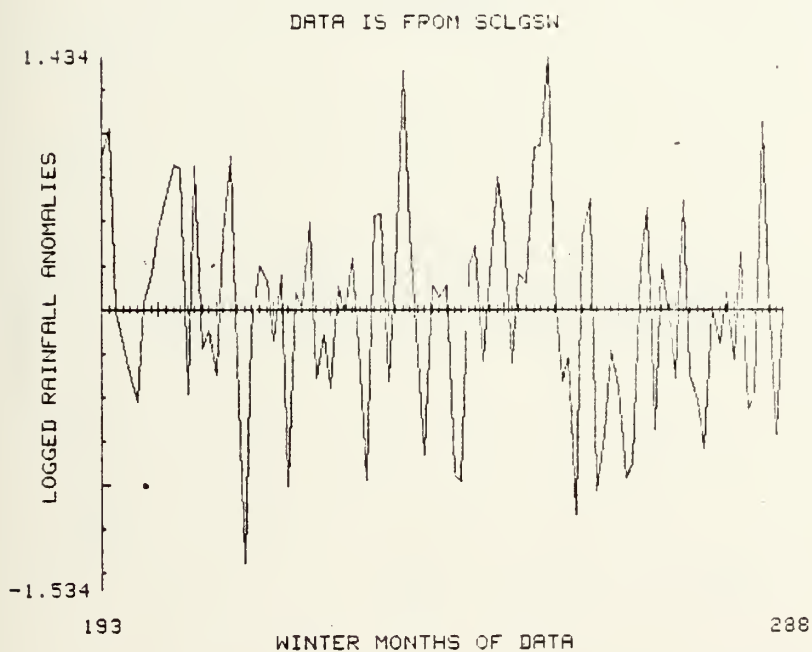
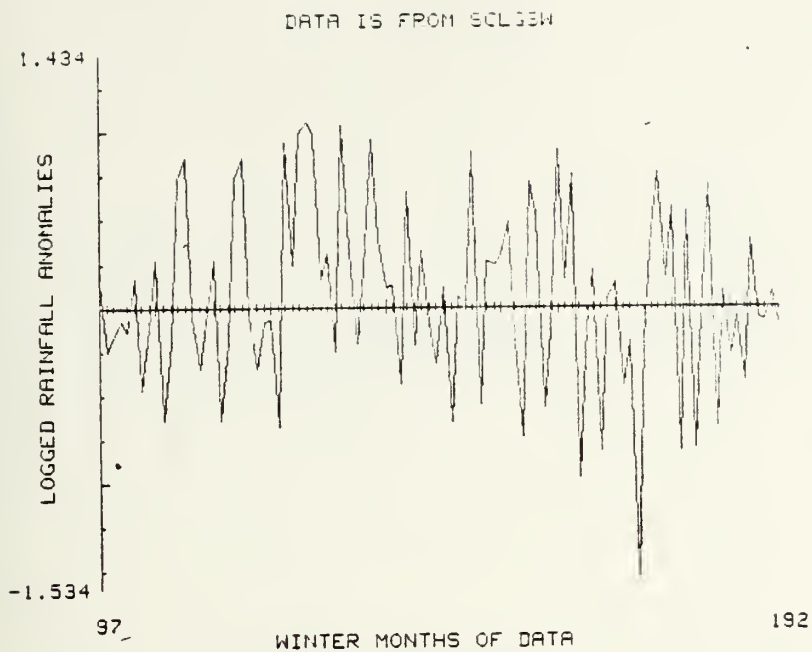


Figure 75b. Years 13 - 36 of winter months only, logged rainfall anomalies of data set SC



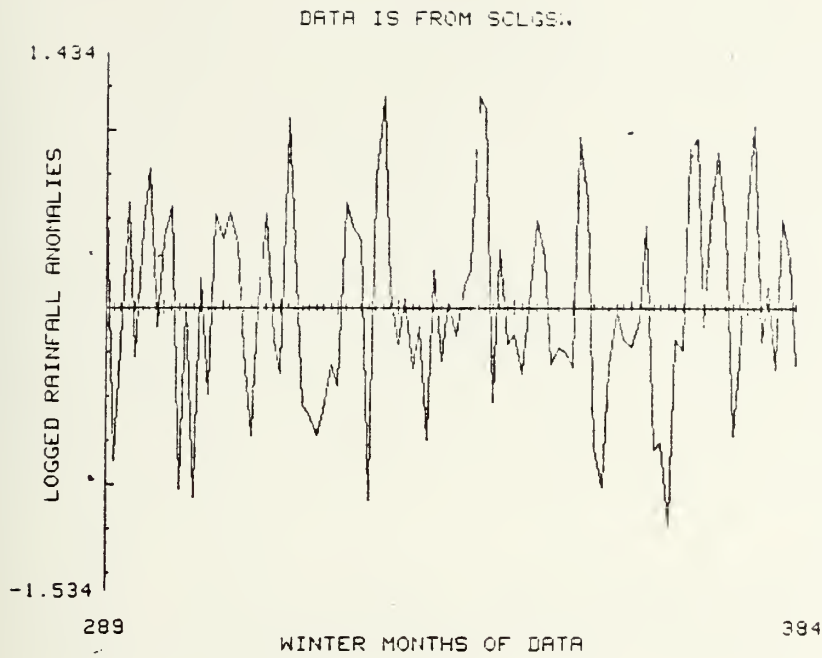


Figure 75c. Years 37 - 48 of winter months only, logged rainfall anomalies of data set SC

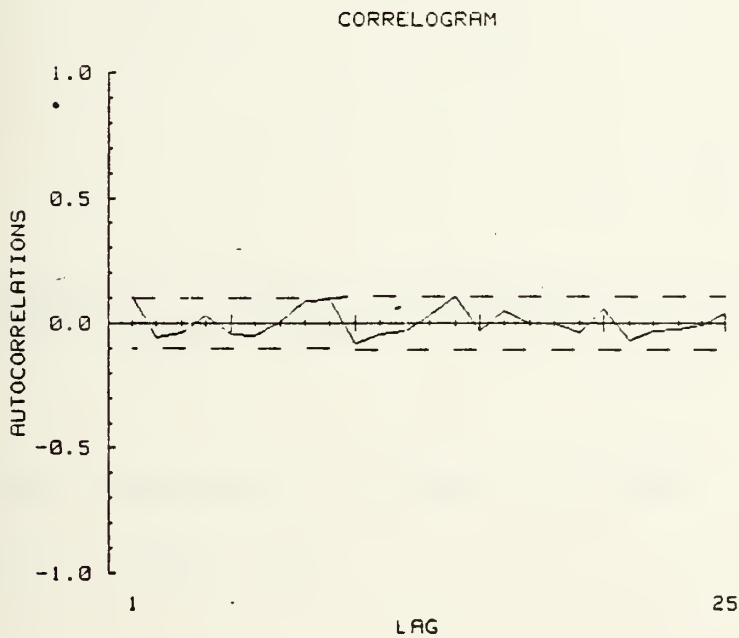


Figure 76. Correlogram of winter months only, logged rainfall anomalies from data set SC





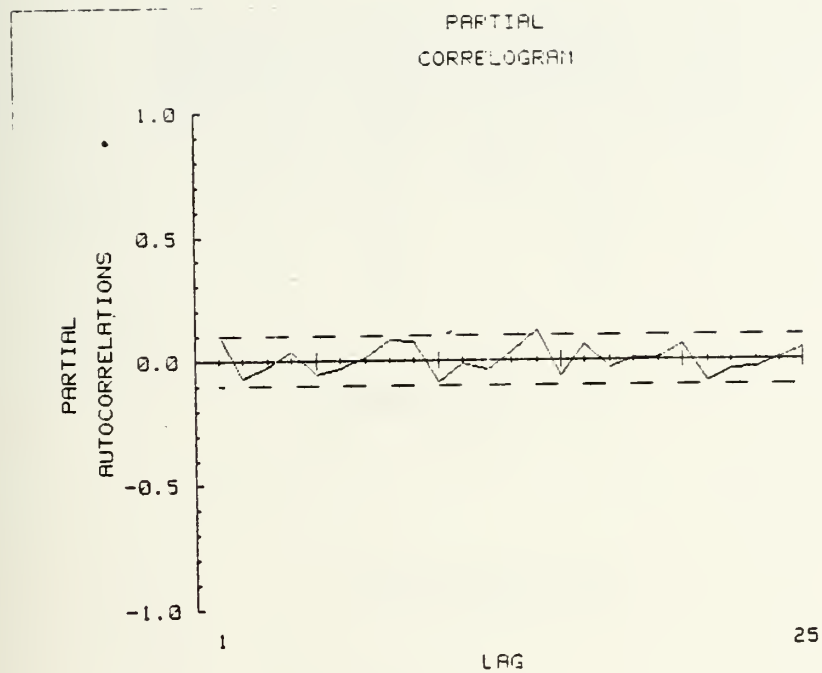


Figure 77. Partial correlogram of winter months only, logged rainfall anomalies from data set SC

This information indicates the model

$$\tilde{z}_t = .107\tilde{z}_t'' + a_t.$$

III.13



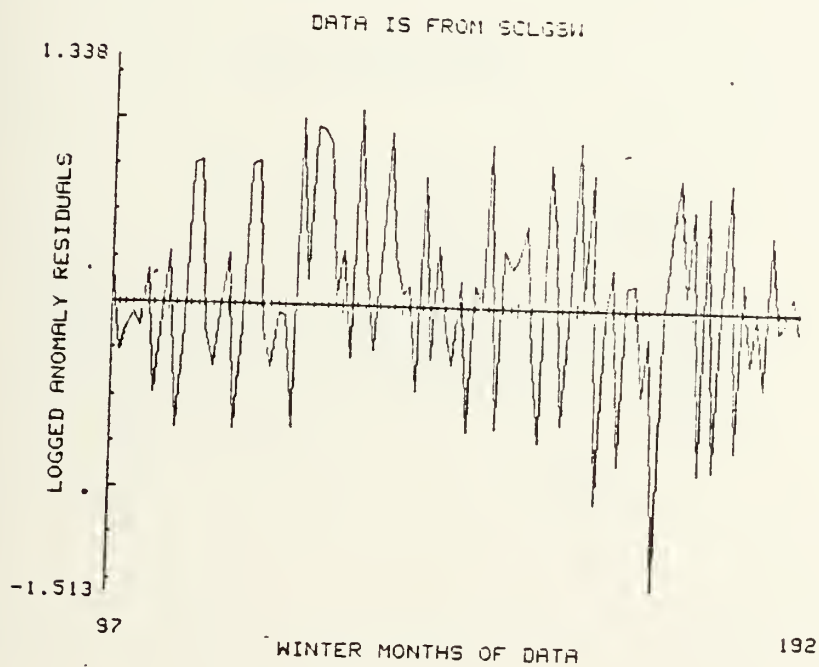
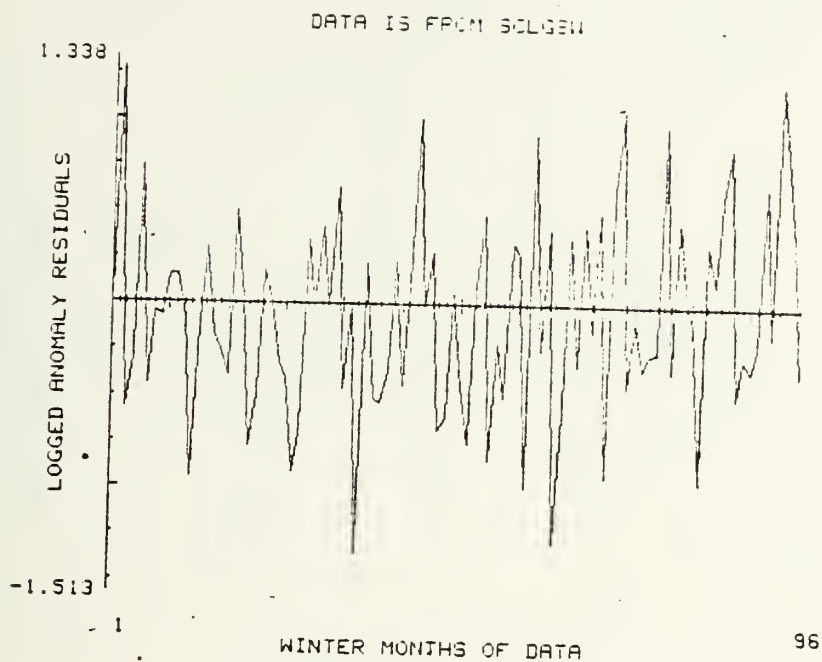


Figure 78a. Years - 24 first order Markov residuals of logged rainfall anomalies, for winter months only, data set SC



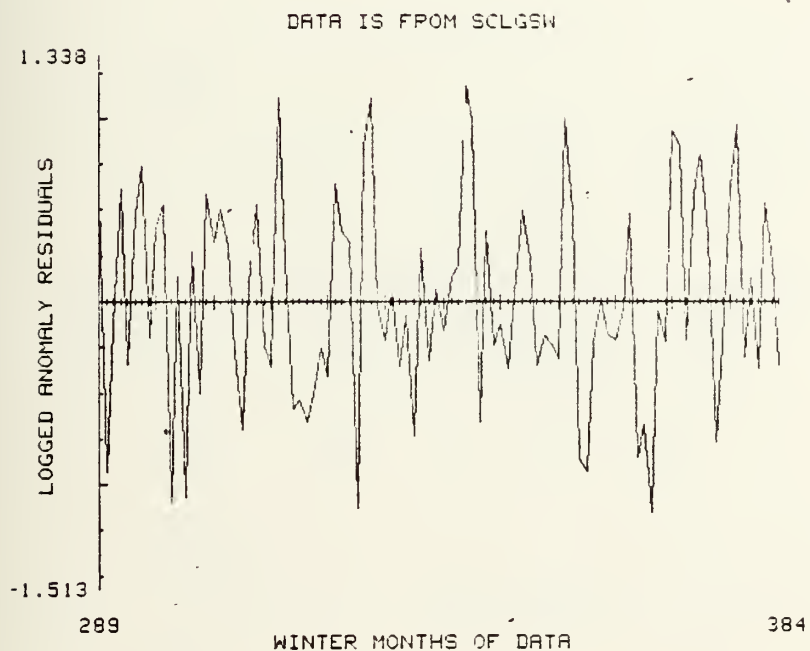
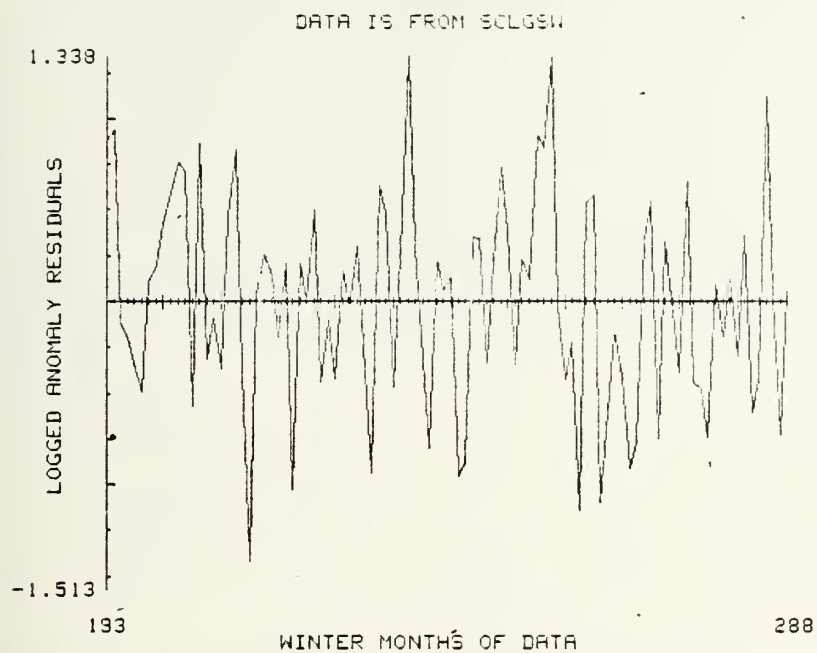


Figure 78b. Years 25 - 48, first order Markov residuals of logged rainfall anomalies for winter months only, data set SC



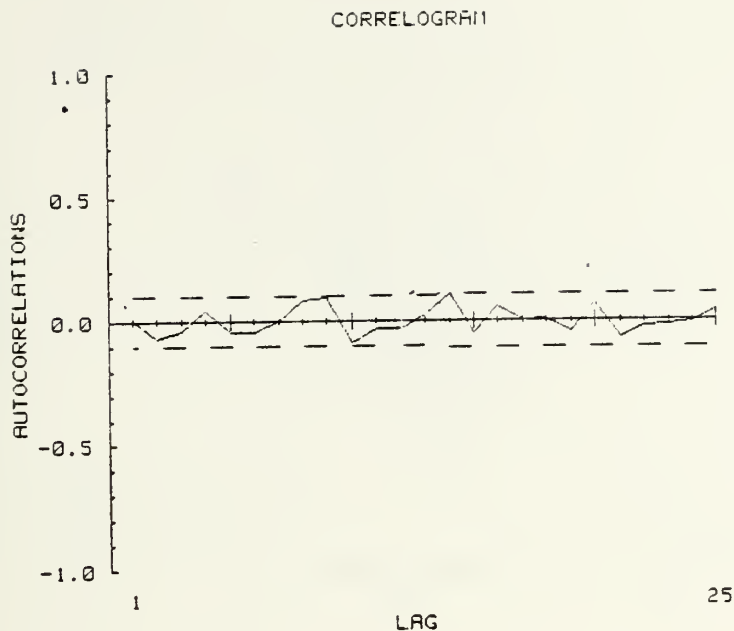


Figure 79. Correlogram of first order Markov residuals of logged rainfall anomalies from winter months only, data set SC.

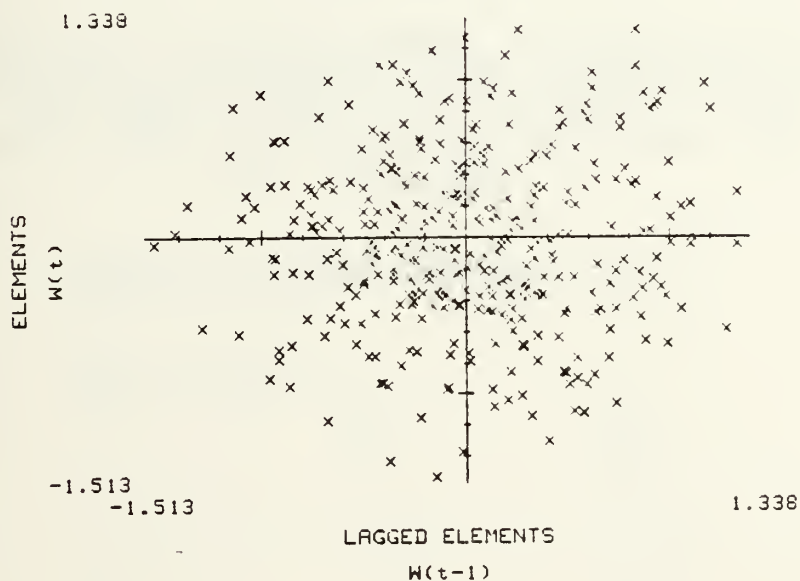


Figure 80. Lag one plot of first order Markov residuals from logged rainfall anomalies of winter months only, data set SC





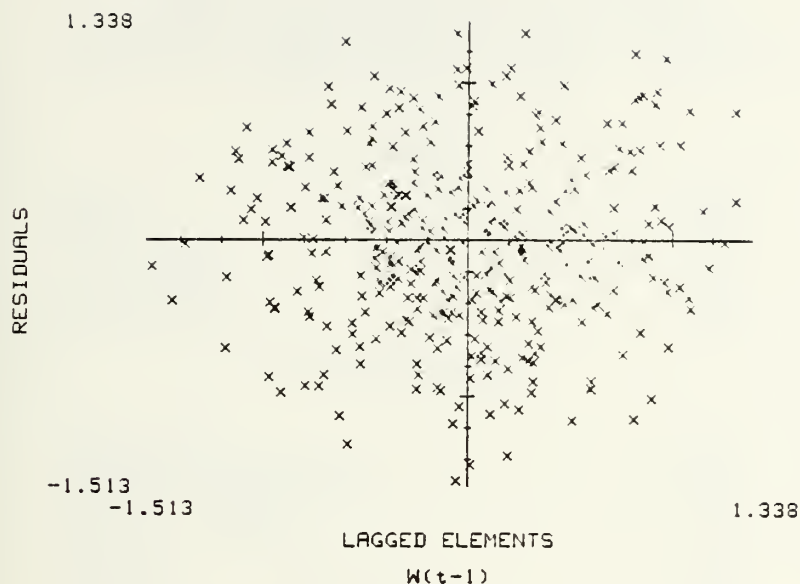


Figure 81. First order Markov residuals versus lag one data point from logged rainfall anomalies of winter months only, data set SC

TABLE 29

ACTUAL AND EXPECTED NUMBER OF TURNING POINTS  
AND ACTUAL AND EXPECTED PHASE FREQUENCIES FROM  
THE FIRST ORDER MARKOV RESIDUALS OF THE  
LOGGED RAINFALL ANOMALIES OF THE WINTER MONTHS  
ONLY, DATA SET SC

NUMBER OF TURNING POINTS = 32  
E[P] = 30.667      V[P] = 15.39

PHASE LENGTHS

| D      | OBS. | E[*] |
|--------|------|------|
| 1      | 21   | 18.8 |
| 2      | 8    | 8.1  |
| 3      | 3    | 2.3  |
| 4      | 0    | .5   |
| 5      | 0    | 0.0  |
| 6      | 0    | 0.0  |
| 7      | 0    | 0.0  |
| 8      | 0    | 0.0  |
| 9      | 0    | 0.0  |
| 10     | 0    | 0.0  |
| TOTALS | 32   | 29.7 |



TABLE 30

GENERAL STATISTICS OF FIRST ORDER  
MARKOV RESIDUALS FROM LOGGED  
RAINFALL ANOMALIES OF WINTER MONTHS  
ONLY, DATA SET SC

Moments

|          |       |
|----------|-------|
| Mean     | .000  |
| Variance | .305  |
| Skewness | .015  |
| Kurtosis | -.363 |

Percentiles

|                 |        |
|-----------------|--------|
| Minimum         | -1.513 |
| Lower Sixteenth | -.872  |
| Lower Eight     | -.663  |
| Lower Quartile  | -.359  |
| Median          | -.026  |
| Upper Quartile  | .355   |
| Upper Eight     | .682   |
| Upper Sixteenth | .882   |
| Maximum         | 1.338  |

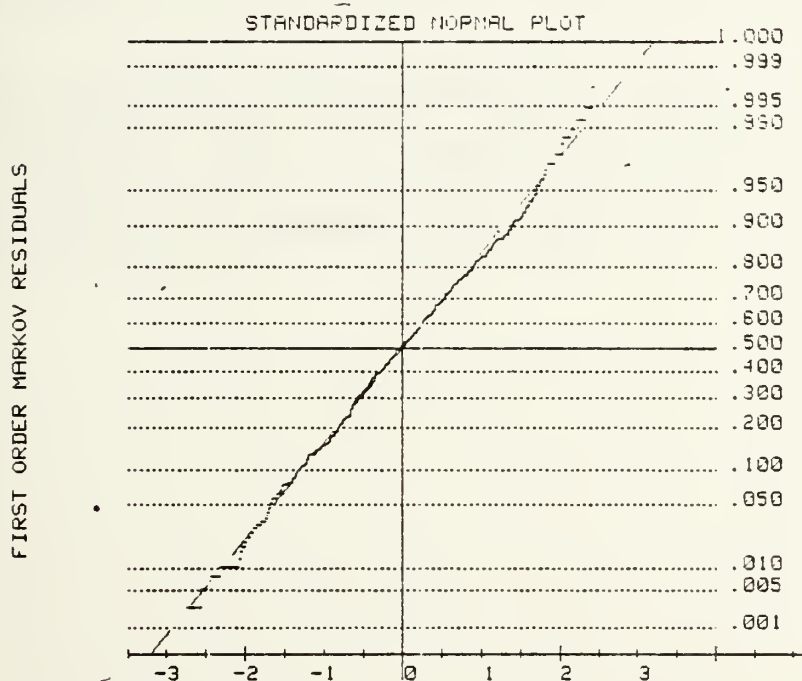


Figure 82. Standardized normal plot of first order Markov residuals from logged rainfall anomalies of winter months only, data set SC



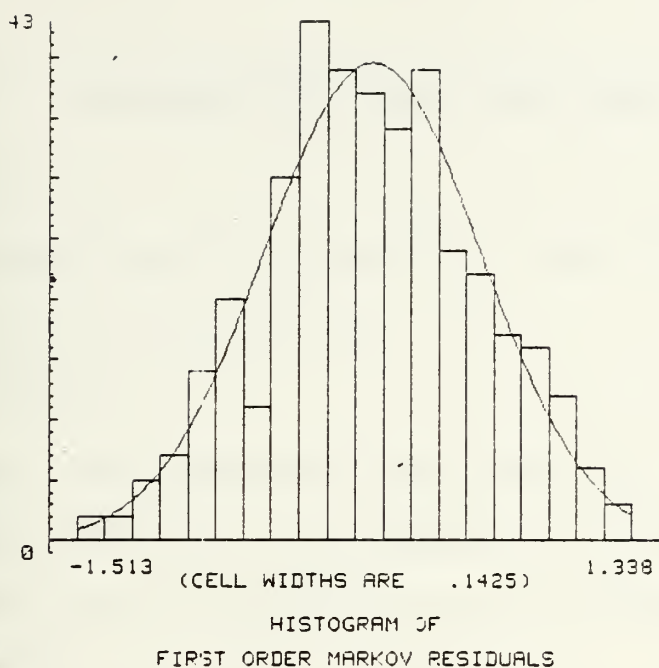


Figure 83. Histogram of first order Markov residuals from logged rainfall anomalies of winter months only, data set SC

The chi-square was calculated at 16.60 for 17 degrees of freedom. This is significant at the .482 level thus indicating probable normality.



#### IV. VALIDATION OF FIRST ORDER MARKOV MODELS

##### A. THEORY

The general model proposed by a first order Markov process is, as stated before;

$$\tilde{z}_t = \rho \tilde{z}_{t-1} + a_t \quad \text{IV.1}$$

where  $\{a_t\}$  are distributed iid  $N(0, \sigma_a^2)$ . To validate this model, preferably independent data should be subjected to the model, and an analysis of the residual, or forecast errors, made.

As stated previously, years 1975 through 1980 were reserved for the purpose of validation. The method of validation was to use the model to construct a series of one step ahead forecasts. Let  $e_t(1)$  be the error in a forecast of time  $t+1$  from the model at time  $t$ . Then the minimum mean squared error forecast (see Box and Jenkins) is;

$$e_t(1) = \tilde{z}_t - \hat{\rho} \tilde{z}_{t-1} \quad \text{IV.2}$$

If the model is correct, the sequence  $\{e_t(1)\}$  will be independent normally distributed with mean zero and variance  $\sigma_a^2$ . In the following sections the models are applied to the reserved data sets (which may also be found in the appendixes), and these forecast errors are calculated. The forecast errors are then analyzed to determine if

- (1) The errors are serially independent





(2) The errors are distributed as normal random variables with mean zero, and variance  $\sigma_a^2$ .

Since the residual analysis of the twelve month model already indicates a poor fit, the twelve month model will not be validated. Only the winter month models will be checked for validity.

#### B. DATA SET RN

Figures 84 (Raw data) and 85 (Logged anomalies) display the reserved data set. The logged anomalies were formed by removing the means of the analyzed data, Table 4, not the means of the logged reserved data. This was done to remove any bias from the validation.

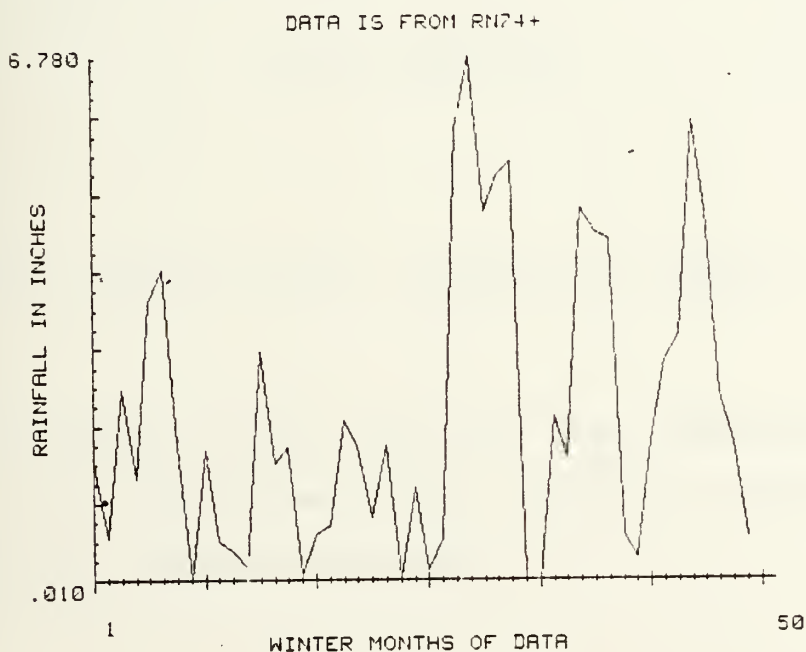


Figure 84. Reserved rainfall data for data set RN



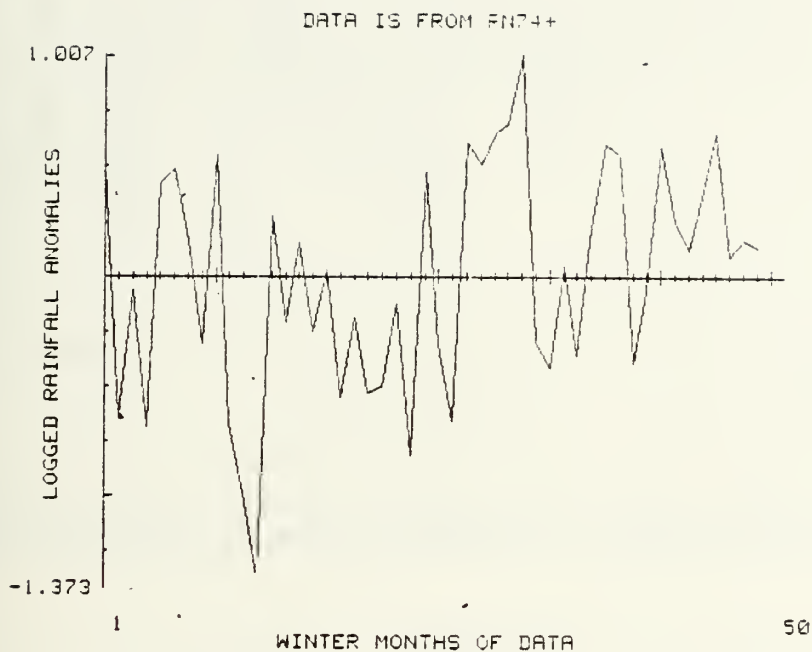


Figure 85. Logged rainfall anomalies of reserved data set RN

The forecast errors, Figure 86, their correlogram, Figure 87, and independence tests, Table 31 indicate that the errors are indeed independent.



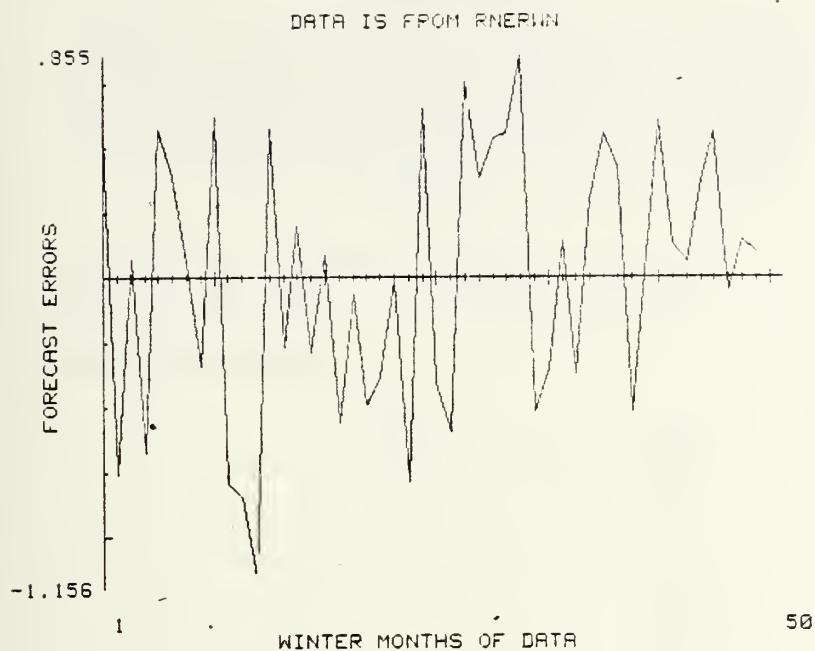


Figure 86. Forecast errors from first order Markov model applied to winter months of reserved data set RN

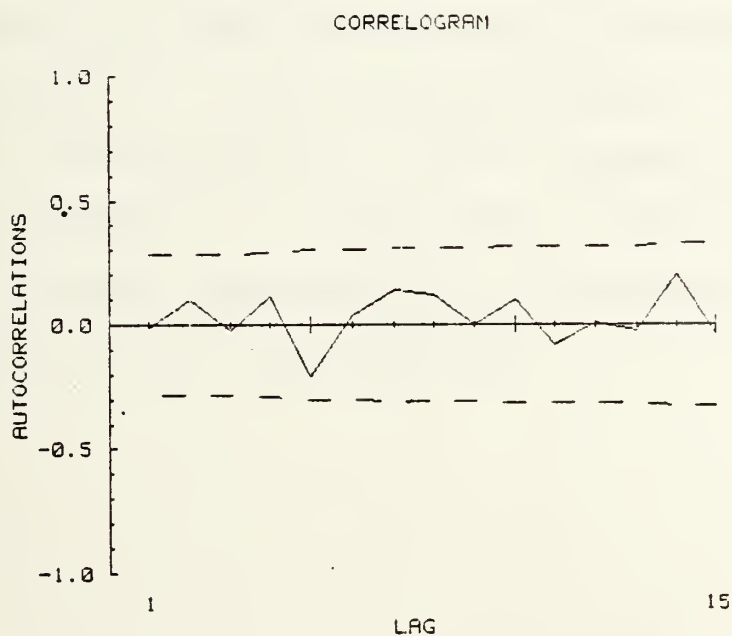


Figure 87. Correlogram of forecast errors from first order Markov model applied to winter months of reserved data set RN



TABLE 31

ACTUAL AND EXPECTED NUMBER OF TURNING  
POINTS AND ACTUAL AND EXPECTED PHASE  
FREQUENCIES FOR THE FORECAST ERRORS OF  
THE FIRST ORDER MARKOV MODEL APPLIED  
TO THE WINTER MONTHS OF RESERVED DATA  
SET RN

NUMBER OF TURNING POINTS = 32  
 $E[P] = 30.667$                        $V[P] = 15.39$

## PHASE LENGTHS

| D      | OBS. | $E[*]$ |
|--------|------|--------|
| 1      | 21   | 18.8   |
| 2      | 8    | 8.1    |
| 3      | 3    | 2.3    |
| 4      | 0    | .5     |
| 5      | 0    | 0.0    |
| 6      | 0    | 0.0    |
| 7      | 0    | 0.0    |
| 8      | 0    | 0.0    |
| 9      | 0    | 0.0    |
| 10     | 0    | 0.0    |
| TOTALS | 32   | 29.7   |

The normality of the forecast errors is addressed by Table 32 (Statistics), Figure 88 (Normal plot), 89 (Histogram), and a simple chi-squared test. The chi-squared was calculated as 7.82 with 5 degrees of freedom which is significant at the .167 level. However, the normality of the errors is somewhat questionable due to the other displays.





TABLE 32

GENERAL STATISTICS OF FORECAST ERRORS FROM  
THE FIRST ORDER MARKOV MODEL APPLIED TO THE  
WINTER MONTHS OF RESERVED DATA SET RN

## Moments

|          |       |
|----------|-------|
| Mean     | .001  |
| Variance | .259  |
| Skewness | -.276 |
| Kurtosis | -.955 |

## Percentiles

|                 |        |
|-----------------|--------|
| Minimum         | -1.156 |
| Lower Sixteenth | -.791  |
| Lower Eight     | -.634  |
| Lower Quartile  | -.397  |
| Median          | .069   |
| Upper Quartile  | .409   |
| Upper Eight     | .569   |
| Upper Sixteenth | .614   |
| Maximum         | .855   |

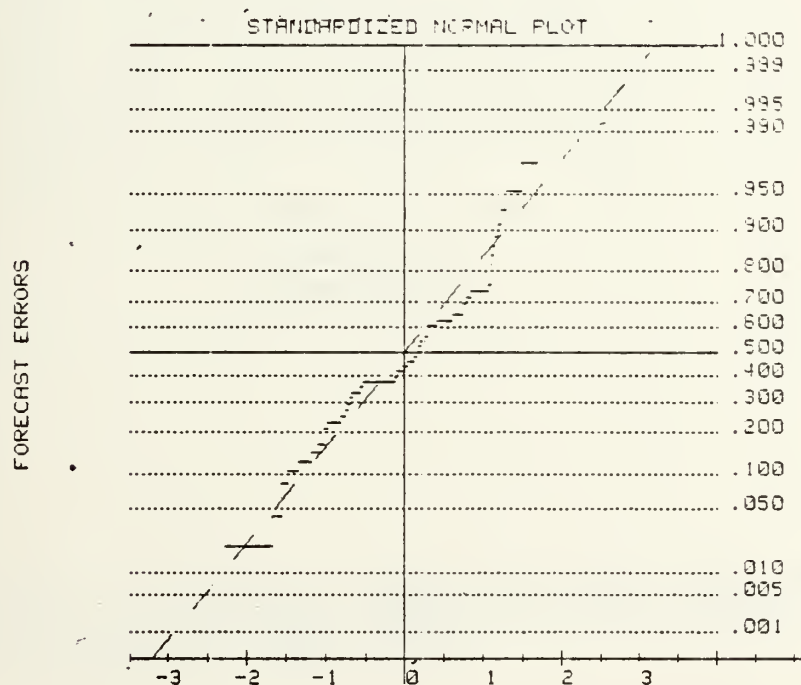


Figure 88. Standardized normal plot of forecast errors from the first order Markov model applied to the winter months of reserved data set RN



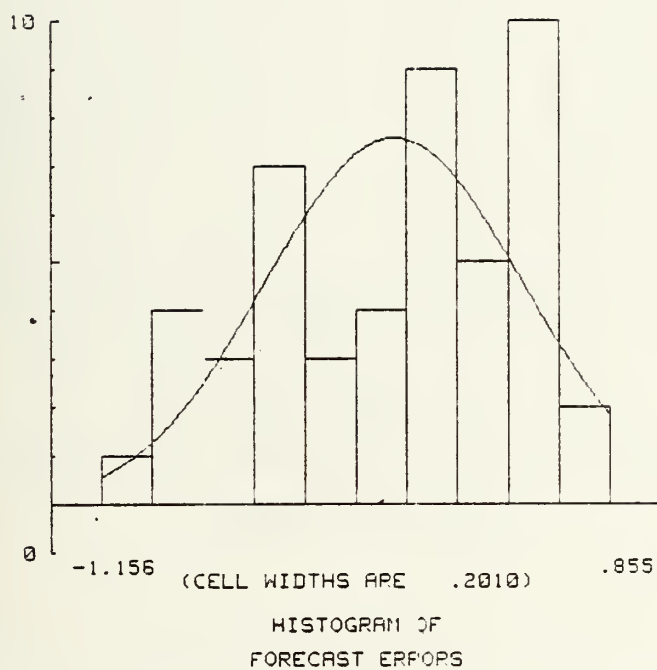


Figure 89. Histogram of forecast errors from the first order Markov model applied to the winter months of reserved data set RN



C. DATA SET FL

As before, the similarity of results for the different data set allows the analysis to be portrayed using the displays only.

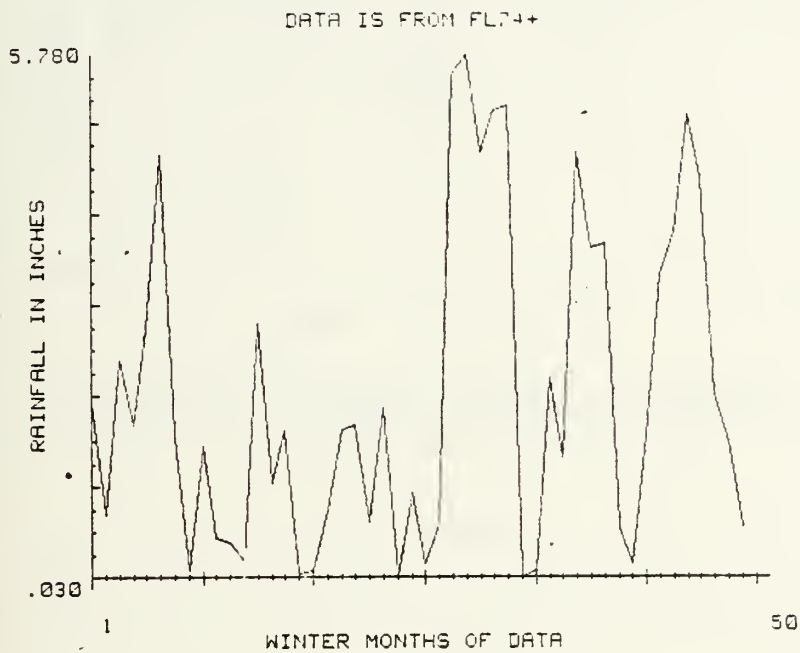


Figure 90. Reserved rainfall data for data set FL



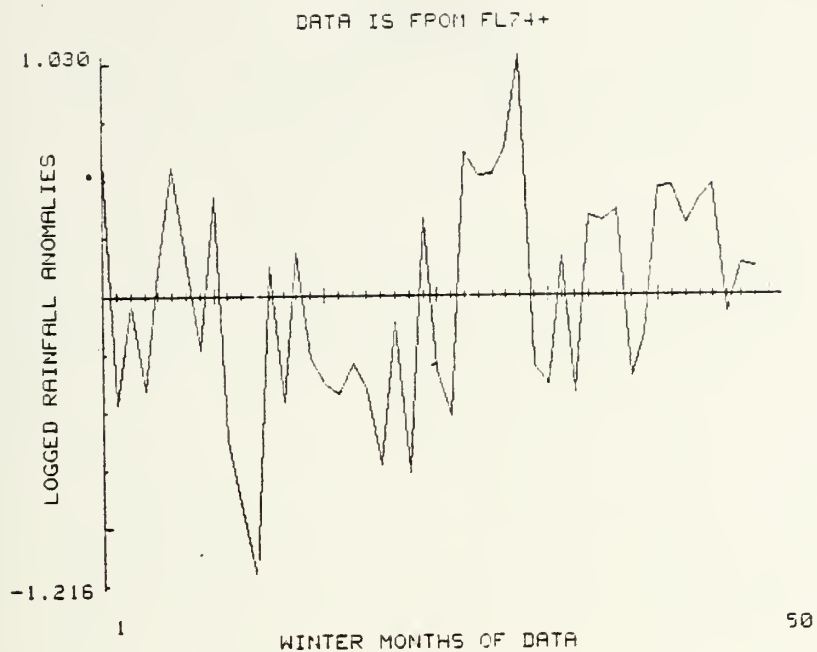


Figure 91. Logged rainfall anomalies of reserved data set FL.

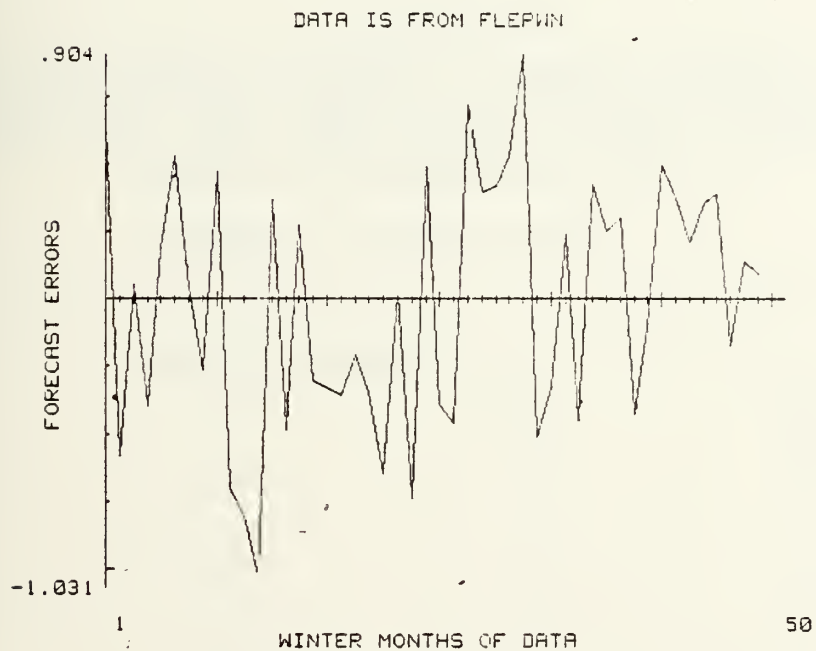


Figure 92. Forecast errors from the first order Markov model applied to the winter months of reserved data set FL





# CORRELOGRAM

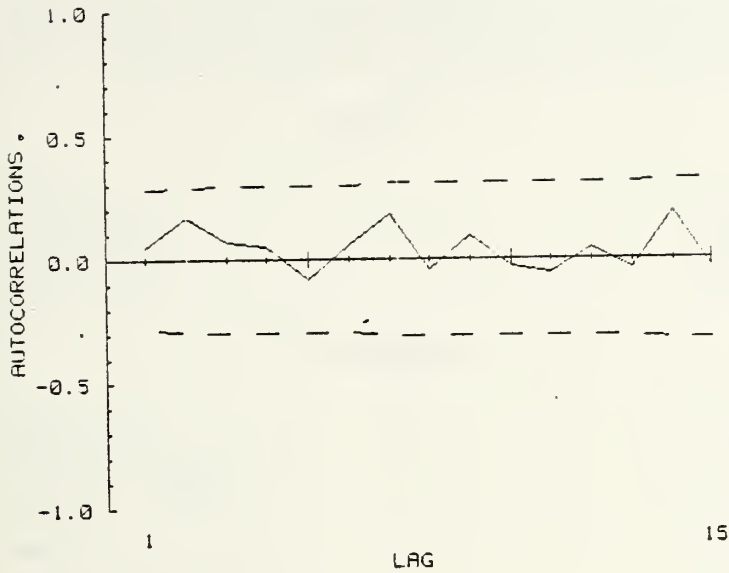


Figure 93. Correlogram of forecast errors from first order Markov model applied to the winter months of reserved data set FL

TABLE 33

ACTUAL AND EXPECTED NUMBER OF TURNING POINTS AND ACTUAL AND EXPECTED PHASE FREQUENCIES FROM THE FORECAST ERRORS OF THE FIRST ORDER MARKOV MODEL APPLIED TO THE WINTER MONTHS OF RESERVED DATA SET FL

NUMBER OF TURNING POINTS = 34

$E[P] = 30.667$

$V[P] = 15.396$

## PHASE LENGTHS

| D      | OBS. | $E[*]$ |
|--------|------|--------|
| 1      | 24   | 18.8   |
| 2      | 8    | 8.1    |
| 3      | 2    | 2.3    |
| 4      | 0    | .5     |
| 5      | 0    | 0.0    |
| 6      | 0    | 0.0    |
| 7      | 0    | 0.0    |
| 8      | 0    | 0.0    |
| 9      | 0    | 0.0    |
| 10     | 0    | 0.0    |
| TOTALS | 34   | 29.7   |



TABLE 34

GENERAL STATISTICS OF FORECAST ERRORS  
FROM THE FIRST ORDER MARKOV MODEL APPLIED  
TO THE WINTER MONTHS OF RESERVED DATA SET FL

## Moments

|          |       |
|----------|-------|
| Mean     | -.015 |
| Variance | .215  |
| Skewness | -.155 |
| Kurtosis | -.970 |

## Percentiles

|                 |        |
|-----------------|--------|
| Minimum         | -1.031 |
| Lower Sixteenth | -.724  |
| Lower Eight     | -.549  |
| Lower Quartile  | -.396  |
| Median          | .049   |
| Upper Quartile  | .369   |
| Upper Eight     | .479   |
| Upper Sixteenth | .531   |
| Maximum         | .904   |

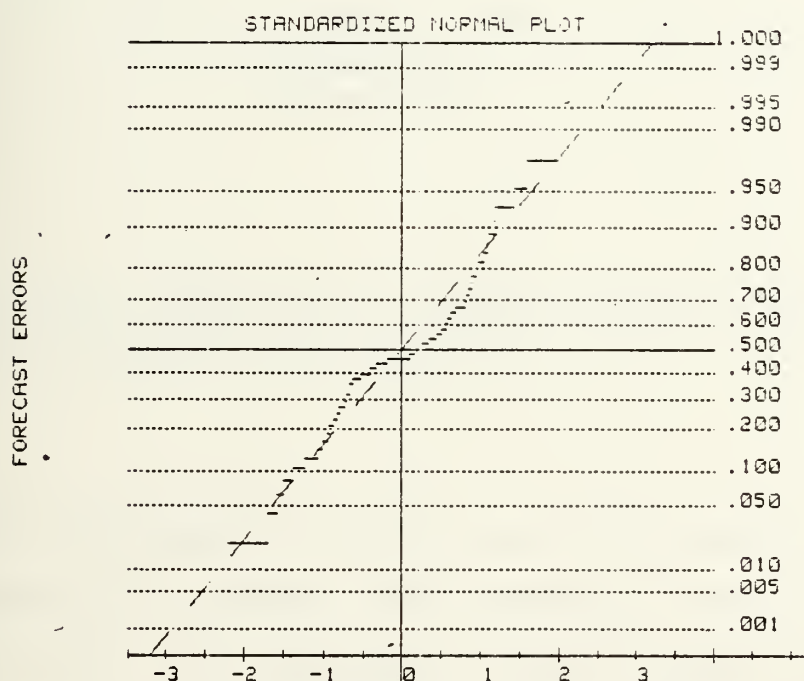


Figure 94. Standardized normal plot of forecast errors from the first order Markov model applied to the winter months of reserved data set FL



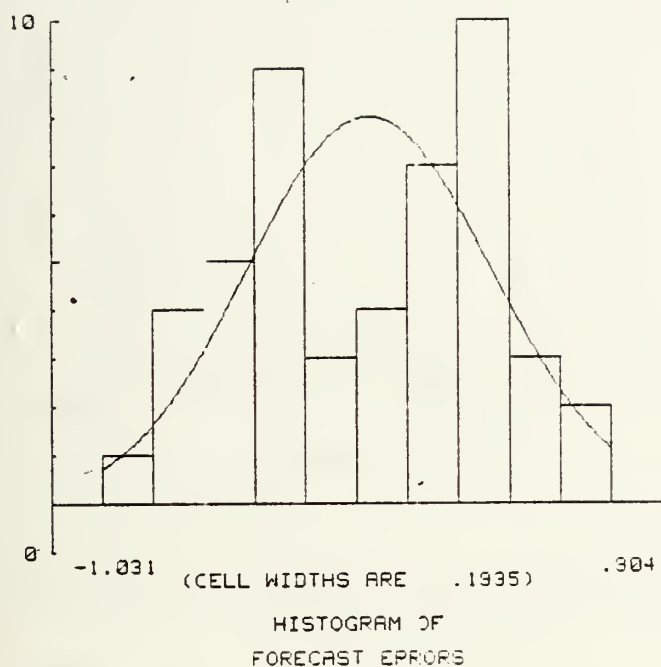


Figure 95. Histogram of forecast errors from the first order Markov model applied to the winter months of reserved data set FL

The chi-squared statistic was calculated as 12.58 with 7 degrees of freedom, thus yielding a significance level of 0.083. This statistic and the displays imply that the data are only marginally normal if at all.



D. DATA SET SC

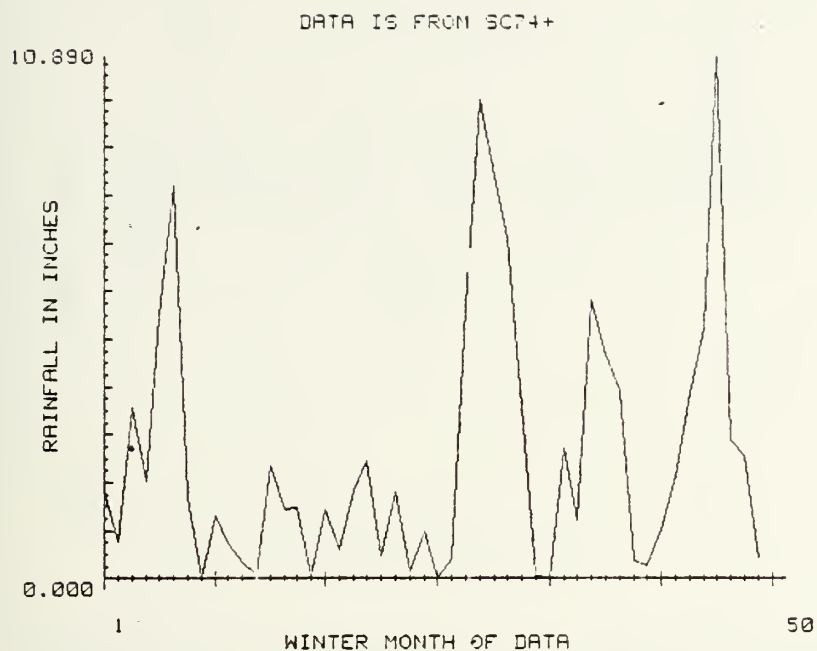


Figure 96. Reserved rainfall for data set SC

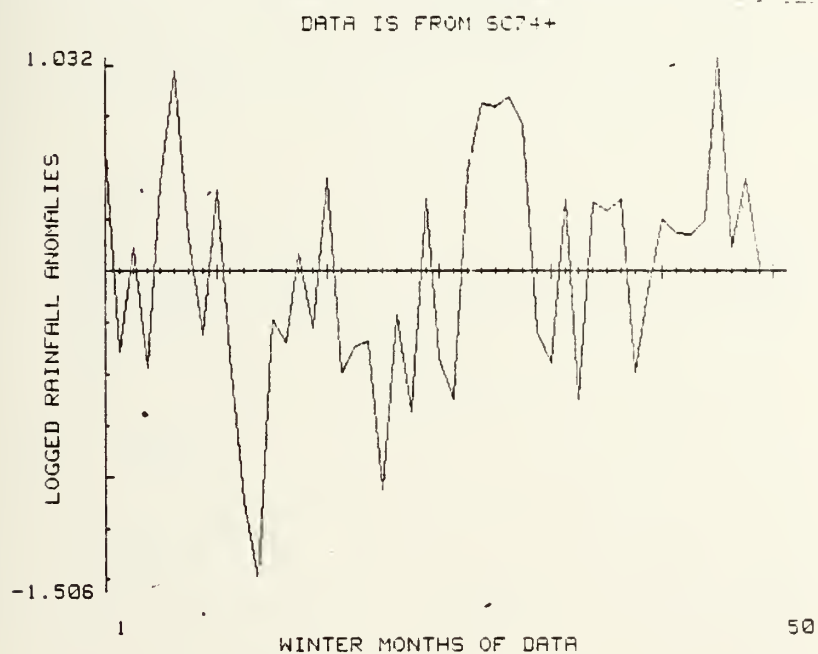


Figure 97. Logged anomalies of reserved data set SC





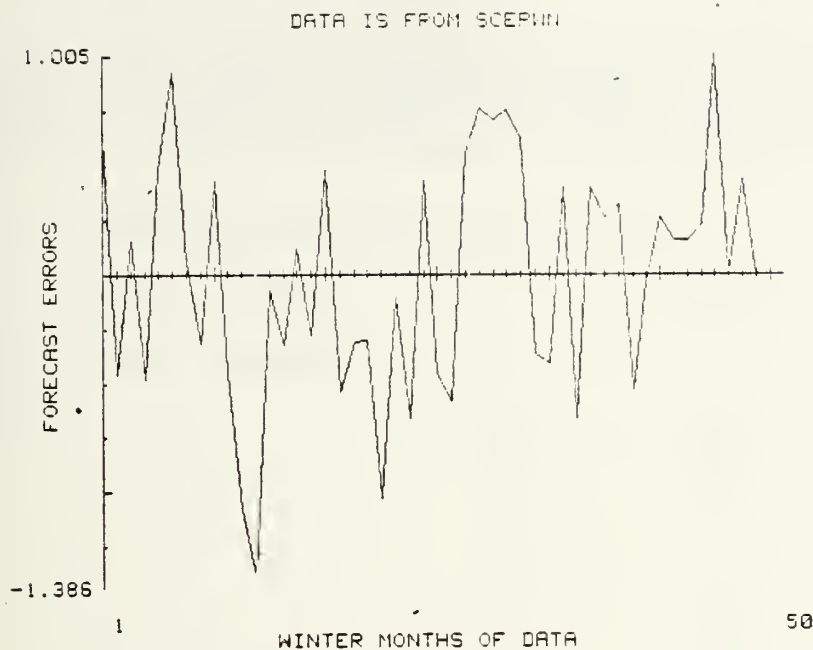


Figure 98. Forecast errors from first order Markov model applied to the winter months of reserved data set SC

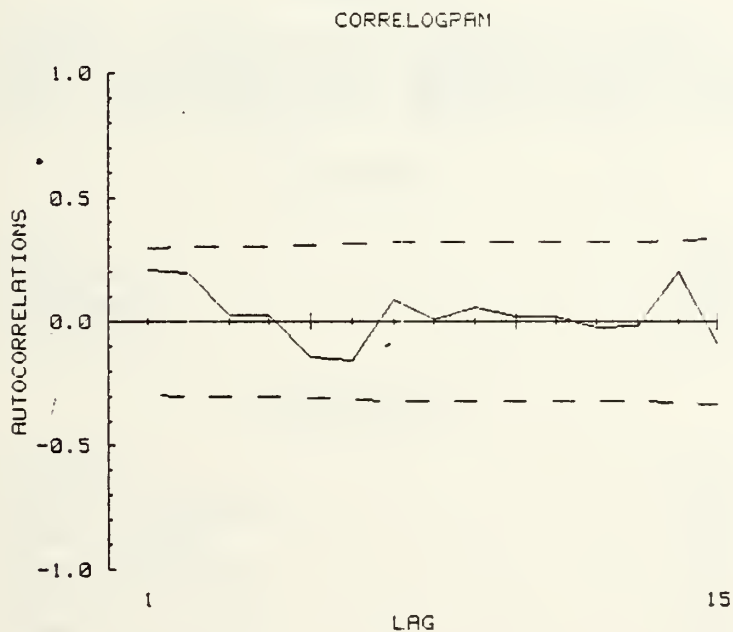


Figure 99. Correlogram of forecast errors from first order Markov model applied to the winter months of reserved data set SC



TABLE 35

ACTUAL AND EXPECTED NUMBER OF TURNING POINTS AND  
ACTUAL AND EXPECTED PHASE FREQUENCES FROM THE  
FORECAST ERRORS OF THE FIRST ORDER MARKOV MODEL  
APPLIED TO THE WINTER MONTHS OF RESERVED DATA SET SC

NUMBER OF TURNING POINTS = 34

$E[P] = 30.667$

$V[P] = 15.396$

PHASE LENGTHS

| D      | OBS. | $E[*]$ |
|--------|------|--------|
| 1      | 24   | 18.8   |
| 2      | 8    | 8.1    |
| 3      | 2    | 2.3    |
| 4      | 0    | .5     |
| 5      | 0    | 0.0    |
| 6      | 0    | 0.0    |
| 7      | 0    | 0.0    |
| TOTALS | 34   | 29.7   |

TABLE 36

GENERAL STATISTICS OF FORECAST ERRORS  
FROM THE FIRST ORDER MARKOV MODEL APPLIED  
TO THE WINTER MONTHS OF RESERVED DATA SET SC

#### Moments

|          |       |
|----------|-------|
| Mean     | -.003 |
| Variance | .296  |
| Skewness | -.298 |
| Kurtosis | -.413 |

#### Percentiles

|                 |        |
|-----------------|--------|
| Minimum         | -1.386 |
| Lower Sixteenth | -.846  |
| Lower Eight     | -.559  |
| Lower Quartile  | -.434  |
| Median          | .031   |
| Upper Quartile  | .414   |
| Upper Eight     | .588   |
| Upper Sixteenth | .735   |
| Maximum         | 1.005  |



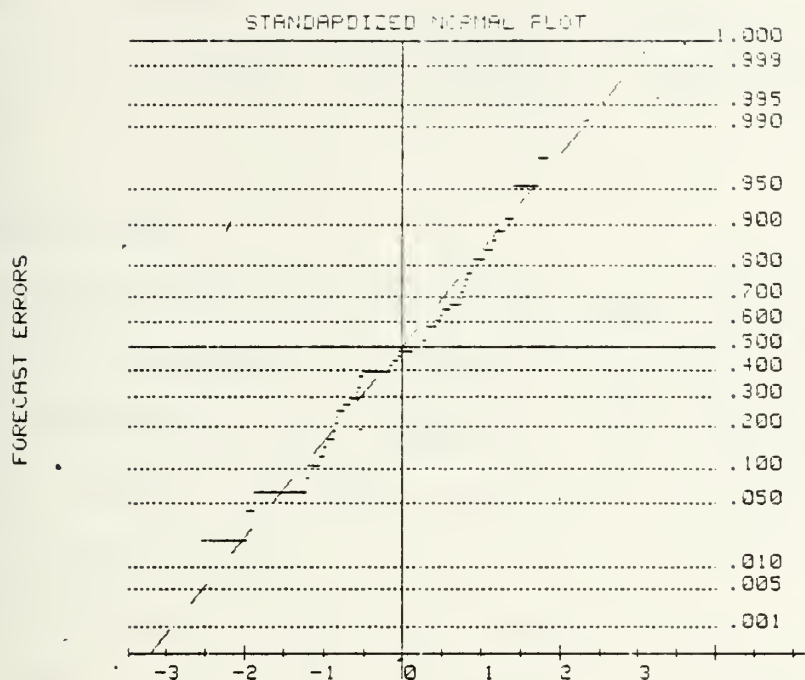


Figure 100. Standardized normal plot of forecast errors from the first order Markov model applied to the winter months of reserved data set SC

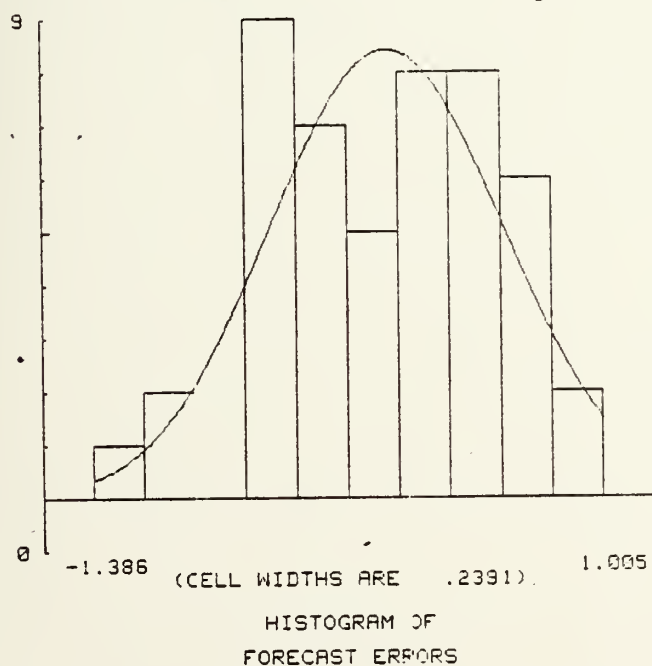


Figure 101. Histogram of forecast errors from the first order Markov model applied to the winter months of reserved data set SC



The chi-squared statistic was calculated as 9.71 with 7 degrees of freedom, thus yielding a significance level of .205.

#### E. CONCLUSIONS

The application of a Markovian model was indicated by the apparent dependence of adjacent months and the apparent lack of dependence at any other lag. The preceeding subsections, however, indicate that the first order Markovian model is weak at best.

The structure of the data, visually, still points toward some sort of underlying order. The following sections attempt to discover this order.





## V. 2x2 TABLES

### A. THEORY

As seen in sections III and IV, the classical ARMA time series approach does not seem to adequately describe the data. Another technique used to explore possible relationships is the 2x2 contingency tables.

The idea to be explored is whether or not some subset of the data, to be called the control, may be used to predict in some way the behavior of another subset of the data, to be called the complement. Here, the data are reduced from monthly observations to yearly observations as described below.

Let  $\underline{X}$  be the subset of a year, to be called the control, and let  $\underline{Y}$  be the subset to be called the complement. It is necessary that  $\underline{X} \cap \underline{Y} = \emptyset$ ; that is, the intersection of these two sets is empty. The data are then compared for some quality in  $\underline{X}$  and for some quality in  $\underline{Y}$ . The question is then: does the presence (or absence) of the quality in  $\underline{X}$  affect the presence (or absence) of the quality in  $\underline{Y}$ ? An example of a typical table is shown below in Figure 102.



|                                 |          | COMPLEMENT |               |               |                  |
|---------------------------------|----------|------------|---------------|---------------|------------------|
|                                 |          | <u>Y</u>   |               |               |                  |
|                                 |          | 1          | 2             |               |                  |
| C<br>O<br>N<br>T<br>R<br>O<br>L | <u>X</u> | 1          | $n_{11}$      | $n_{12}$      | $n_{1\cdot}$     |
|                                 |          | 2          | $n_{21}$      | $n_{22}$      | $n_{2\cdot}$     |
|                                 |          |            | $n_{\cdot 1}$ | $n_{\cdot 2}$ | $n_{\cdot\cdot}$ |

Figure 102. Typical 2x2 contingency table

The table elements,  $n_{ij}$ , represents the number of years which display quality  $i$  in the control and quality  $j$  in the complement. The marginal entries  $n_{i\cdot}$  and  $n_{\cdot j}$  represents the numbers of years for which the control has quality  $i$  and the number of years the complement has quality  $j$  respectively. The overall number of years,  $n_{\cdot\cdot}$ , is in the lower right of the table.

Brownlee [Ref. 6] contains a very good discussion of the theory and use of 2x2 contingency tables. Using the notation of Brownlee, let  $\theta_{ij}$  be the probability that any given year



will have a control quality  $i$  and a complement quality  $j$ . Then estimates of the  $\theta_{ij}$ 's are

$$\begin{aligned}\hat{\theta}_{ij} &= n_{ij}/n_{..} \\ \hat{\theta}_{i.} &= n_{i.}/n_{..} \\ \hat{\theta}_{.j} &= n_{.j}/n\end{aligned}\tag{V.1}$$

If the control and complement are independent,

$$\theta_{ij} = \theta_{i.}\theta_{.j}\tag{V.2}$$

These simple assumptions allow for a thorough investigation of the possible interrelationships within the data sets.

Another way to view the assumption of independence is through the use of proportions. Thus, if the basic division is made via the quality of the control, the the proportions

$$P_1 = n_{11}/n_{1.} \quad (\text{respectively } P_2 = n_{21}/n_{2.})\tag{V.3}$$

represent, in words, the proportion of the years that have quality 1 in the control and have quality 1 (respectively 2) in the complement.

The question of independence may be approached in several ways as described below.

### 1. Fishers Exact Test

A test for the significance of any dependence was proposed by Fisher in the case in which the marginal totals,  $n_{i.}$ ,  $n_{.j}$ , and  $n_{..}$  are known a priori (cf. Brownlee [Ref. 6]). To draw from Brownlee, knowledge of the marginals and  $n_{11}$  gives knowledge of all the other elements of the table. The



probability of the event of having exactly  $n_{11}$  years that display quality 1 in both the complement and the control is;

$$P(N_{11}=n_{11}) = \frac{n_{1.}!n_{.1}!n_{.2}!n_{2.}!}{n_{..}!n_{11}!n_{12}!n_{21}!n_{22}!} \quad \text{V.4}$$

A test may then be applied, using V.4, to determine the significance of any dependence. This test is usually applied by simply summing these probabilities in the tail of the distribution (V.4) in the same direction as the noted extreme.

The usual procedure to provide a two-sided test of significance is to double a one sided figure. This procedure is acceptable due to the symmetric appearance of the distribution.

Under the assumptions of independence

$$E \left[ N_{11} \right] = \frac{n_{1.} \cdot n_{.1}}{n_{..}} \quad \text{V.5}$$

$$V \left[ N_{11} \right] = \frac{n_{1.} \cdot n_{.1} \cdot n_{2.} \cdot n_{.2}}{n_{..} (n_{..} - 1)} \quad \text{V.6}$$

and the random variable  $U$  defined as

$$U = \frac{N_{11} - E[N_{11}]}{\sqrt{V[N_{11}]}} \quad \text{V.7}$$

is asymptotically distributed as a normal random variable with mean zero and variance one. This asymptotic result combined with a continuity correction yields a test statistic





of

$$u' = \frac{\{ |n_{11} \ n_{22} \ -n_{12} \ n_{21} \ | -n_{..}/2 \} \sqrt{n_{..}}}{\sqrt{n_{1.} \cdot n_{.1} \cdot n_{2.} \cdot n_{.2}}} \quad \text{V.8}$$

The statistic  $U'$  may then be used as a test, using standard normal tables, of the significance of any variation from the assumption of independence. It should be noted at this point that if the random variable  $U'$  is squared,  $U'^2$  will be distributed as a chi-square with one degree of freedom. The squaring of  $U'$  with simplifying algebra yields the Yates correction to a standard chi-squared goodness of fit statistic

$$(u')^2 = \frac{\{ |n_{11} \ n_{22} \ -n_{12} \ n_{21} \ | -n_{..}/2 \}^2 \ n_{..}}{(n_{11}+n_{12})(n_{11}+n_{21})(n_{12}+n_{22})(n_{21}+n_{22})} \quad \text{V.9}$$

see Dixon and Massey [Ref. 5]. This allows the use of the chi-square tables as an equivalent test to that of V.8.

## 2. Odds

Subsection V.A.1 above deals with the significance of any observed interdependence between the control and the complement. The question of the degree of dependence should also be addressed. The measure to be used is the odds ratio. Using the notation of Fleiss [Ref. 3], a measure of seeing quality 1 in the complement  $\underline{Y}$  may be

$$\Omega_1 = \frac{P(Y=1 \mid X=1)}{P(Y=2 \mid X=1)} ; \quad \text{V.10}$$

this is then the odds that quality 1 will occur in the complement given that quality 1 is present in the control. In a



similar manner;

$$\Omega_2 = \frac{P(Y=1 | X=2)}{P(Y=2 | X=2)} \quad V.11$$

is the odds that quality 1 will occur in the complement given that quality 2 was observed in the control. The currently most often used measure is the odds ratio  $\omega$ , or

$$\omega = \frac{\Omega_1}{\Omega_2} \quad V.12$$

Note that, if the appearance of quality 1 in the complement is independent of whether or not it appears in the control, then  $\omega = 1$ . While  $\omega > 1$  implies that the odds of the complement having quality 1, given that quality 1 was observed in the control, are greater than the odds of the complement having quality 2. This would indicate that the control would be some sort of predictor for the complement, relative to the selected qualities.

In the same continuity correcting spirit, as was used with the Yates chi-square, an estimate for  $\omega$  may be obtained from the Table as

$$\hat{\omega} = 0 = \frac{(n_{11}+.5)(n_{22}+.5)}{(n_{12}+.5)(n_{21}+.5)} \quad V.13$$

with a standard error of

$$s.e.(0) = 0 \sqrt{\frac{1}{n_{11}+.5} + \frac{1}{n_{12}+.5} + \frac{1}{n_{21}+.5} + \frac{1}{n_{22}+.5}} \quad V.14$$

The natural logarithm of this odds ratio will be discussed more fully in section VI.



## B. ANALYSIS

The theory of subsection A above is applied to the three data sets as discussed below. The control is typically taken as a monthly anomaly, say October. Here, the quality is taken as either a positive or a negative anomaly. Thus  $X=1$  occurs when the month of October falls below its mean and  $X=2$  occurs when it falls above its mean. The complement consists of the sum of the rainfall for the succeeding eleven months, or in symbols:

$$X_t = R_{t,1} - \bar{R}_t.$$

$$Y_t = \sum_{m=2}^{12} R_{t,m} - \frac{1}{N} \sum_{t=1}^N \left( \sum_{m=2}^{12} R_{t,m} \right). \quad V.15$$

Where it is understood that  $X=1$  when  $X_t < 0$ ,  $X=2$  when  $X_t > 0$  and similarly for  $Y_t$ .

Various control subsets are used; October through September were investigated by themselves as were all adjacent pairs, triples, and four-tuples of months. For an example, consider the spring (April, May, and June) and its complement (July through March). In this case

$$X_t = \sum_{m=7}^9 R_{t,m} - \frac{1}{N} \sum_{t=1}^{N-1} \left( \sum_{m=7}^9 R_{t,m} \right) \quad V.16$$

$$Y_t = \sum_{m=10}^{12} R_{t,m} + \sum_{m=1}^6 R_{t+1,m} - \frac{1}{N-1} \sum_{t=1}^{N-1} \left( \sum_{m=10}^{12} R_{t,m} + \sum_{m=1}^6 R_{t+1,m} \right).$$

Equations V.15 and V.16 imply that the data are always analyzed as deviations from the arithmetic mean. However, the data are also analyzed as deviations from the median



and the lower quartile. In the tables to follow, 'A' refers to both control and complement having the arithmetic mean removed, 'M' refers to both control and complement having their respective medians removed, and 'QL' refers to the control having the lower quartile removed while the median was removed from the complement.

The first four Tables (37 through 40), give the significance levels of observed departures from independence of the control and complement. Only those values having a Yates corrected chi-square of greater than 1.00 are listed. The entries represent the two-tailed probability of a random deviation in excess of that observed. Although the cut off criterion was the Yates chi-square, the agreement between its probability and that obtained from the Fisher exact and normal tests did not differ in the first two decimal places.

TABLE 37

SIGNIFICANCE OF OBSERVED DEPARTURES FROM INDEPENDENCE OF  
SINGLE MONTH CONTROL VERSUS SUCCEEDING ELEVEN  
MONTH COMPLEMENTS

| Data set<br>Differentiator | RN  |     |     | FL   |     |     | SC  |     |     |
|----------------------------|-----|-----|-----|------|-----|-----|-----|-----|-----|
|                            | A   | M   | QL  | A    | M   | QL  | A   | M   | QL  |
| Control                    |     |     |     |      |     |     |     |     |     |
| October                    | .18 | .22 | .24 |      |     | .31 |     | .15 |     |
| November                   |     |     |     |      | .10 | .12 |     |     | .28 |
| December                   |     |     |     |      |     |     |     |     |     |
| January                    | .14 | .30 | .19 | .002 | .04 | .12 | .21 | .19 |     |
| February                   |     |     |     |      |     |     |     | .11 |     |
| March                      |     |     |     |      |     |     |     | .13 |     |
| April                      |     |     |     |      |     |     |     | .18 |     |
| May                        |     | .14 |     |      |     |     |     |     |     |
| June                       | .24 |     |     |      |     |     |     |     |     |
| July                       |     | .31 |     |      |     |     | .21 |     |     |
| August                     |     |     |     |      |     |     |     |     |     |
| September                  |     | .14 | .12 |      |     |     |     |     |     |





TABLE 38

SIGNIFICANCE OF OBSERVED DEPARTURES FROM INDEPENDENCE  
OF PAIRS OF MONTHS VERSUS SUCCEEDING  
TEN MONTH COMPLEMENTS

| Data Set        | RN  |     |     | FL  |     |     | SC  |     |     |
|-----------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Differentiators | A   | M   | QL  | A   | M   | QL  | A   | M   | QL  |
| Control         |     |     |     |     |     |     |     |     |     |
| Oct+Nov         | .18 |     |     |     |     | .09 | .24 |     |     |
| Nov+Dec         |     |     |     |     | .10 | .12 | .15 | .31 | .08 |
| Jan+Feb         | .30 | .30 |     | .01 | .02 | .12 | .11 | .19 |     |
| Feb+Mar         |     |     |     |     |     |     |     |     | .11 |
| Mar+Apr         |     |     |     |     |     |     |     |     |     |
| Apr+May         |     |     |     |     |     |     |     |     |     |
| May+Jun         |     | .14 | .12 | .24 |     |     |     |     |     |
| Jun+Jul         |     |     |     |     |     |     |     |     |     |
| Jul+Aug         |     |     | .19 |     | .18 |     |     |     |     |
| Aug+Sep         |     |     |     |     |     |     |     |     |     |

TABLE 39

SIGNIFICANCE OF OBSERVED DEPARTURES FROM INDEPENDENCE  
OF TRIPLES OF MONTHS VERSUS SUCCEEDING  
NINE-MONTH COMPLEMENTS

| Data Set        | RN  |     |     | FL  |     |    | SC  |     |    |
|-----------------|-----|-----|-----|-----|-----|----|-----|-----|----|
| Differentiators | A   | M   | QL  | A   | M   | QL | A   | M   | QL |
| Control         |     |     |     |     |     |    |     |     |    |
| Oct+Nov+Dec     |     |     |     |     |     |    | .19 |     |    |
| Nov+Dec+Jan     |     |     |     |     |     |    |     |     |    |
| Dec+Jan+Feb     |     |     |     | .31 |     |    |     |     |    |
| Jan+Feb+Mar     |     |     |     |     |     |    |     |     |    |
| Feb+Mar+Apr     |     |     |     |     |     |    | .31 | .28 |    |
| Mar+Apr+May     |     |     |     |     |     |    | .22 |     |    |
| Apr+May+Jun     |     |     | .30 |     |     |    |     |     |    |
| May+Jun+Jul     | .14 | .30 |     |     |     |    |     |     |    |
| Jun+Jul+Aug     |     |     |     |     | .18 |    |     |     |    |
| Jul+Aug+Sep     |     |     |     |     |     |    |     |     |    |



TABLE 40

SIGNIFICANCE OF OBSERVED DEPARTURES FROM  
INDEPENDENCE OF FOUR-TUPLES OF MONTHS VERSUS  
SUCCEEDING EIGHT-MONTH COMPLEMENTS

| Data set        | RN  |     |     | FL  |   |    | SC  |     |     |
|-----------------|-----|-----|-----|-----|---|----|-----|-----|-----|
| Differentiator  | A   | M   | QL  | A   | M | QL | A   | M   | QL  |
| Control         |     |     |     |     |   |    |     |     |     |
| Oct+Nov+Dec+Jan |     |     |     | .26 |   |    |     |     |     |
| Nov+Dec+Jan+Feb |     | .22 |     |     |   |    |     |     |     |
| Dec+Jan+Feb+Mar |     | .22 |     |     |   |    |     |     |     |
| Jan+Feb+Mar+Apr |     |     |     |     |   |    |     | .31 |     |
| Feb+Mar+Apr+May | .22 |     |     |     |   |    | .27 |     |     |
| Mar+Apr+May+Jun |     |     |     |     |   |    | .20 | .31 |     |
| Apr+May+Jun+Jul |     |     | .30 |     |   |    |     |     | .28 |
| May+Jun+Jul+Aug |     | .14 |     |     |   |    |     |     |     |
| Jun+Jul+Aug+Sep |     |     |     |     |   |    |     |     |     |

Several choices for predictors are suggested in the previous tables. However, the apparent strongest candidate for a predictor is January. The control of January by itself and January paired with December, are the most consistently significant entries. Tables 41 below gives the odds ratio, V.13, for January, and January and December, as controls.



TABLE 41

ODDS RATIO OF JANUARY VERSUS FEBRUARY  
THROUGH DECEMBER AND JANUARY PLUS  
DECEMBER VERSUS FEBRUARY THROUGH NOVEMBER

| Differentiator | A     | M    | QL   |
|----------------|-------|------|------|
| Data set RN    |       |      |      |
| January        | 4.59  | 3.15 | 5.13 |
| Jan+Dec        | 3.15  | 3.15 | 2.01 |
| Data set FL    |       |      |      |
| January        | 10.71 | 4.72 | 4.30 |
| Jan+Dec        | 7.42  | 6.02 | 4.30 |
| Data set SC    |       |      |      |
| January        | 2.45  | 2.49 | 2.23 |
| Jan+Dec        | 3.00  | 2.49 | 1.44 |

At this point in the analysis it was decided to explore more fully the power of January as a predictor. It should be stated that other possibilities for predictors are suggested by the tables, but time did not allow an exhaustive study of all of these.

#### C. OTHER RESULTS

The results of section V.B suggest that a more detailed analysis of January as a predictor is in order. The first method tried for this was ordinary least squares regression of the rainfall total in January versus the total for February through December. This is the model below.

$$\text{Let } X_t = R_{t,4}$$

$$Y_t = \sum_{m=5}^{12} R_{t,m} + \sum_{m=1}^3 R_{t+1,m} \quad \text{V.17}$$



then assume that

$$Y_t = \alpha + \beta X_t + \epsilon_t \quad V.18$$

as the standard, linear model where  $\{\epsilon_t\}$  are assumed to be independent and identically distributed with mean zero and variance  $\sigma\epsilon^2$ . If the predictability of January is strong, this model, V.18, may result in a good fit of the data. Table 42 below is the resulting ANOVA for this regression. As may easily be seen, the model does not appear to have any significance.

TABLE 42

ANOVA FOR REGRESSION OF SIMPLE  
LINEAR MODEL FOR ALL DATA SETS

R-squared = .1033

Standard error of estimate = 4.1574

RN

| SOURCE      | DF | SS      | MS     | F    |
|-------------|----|---------|--------|------|
| Total       | 22 | 404.798 |        |      |
| Regression  | 1  | 41.821  | 41.821 | 2.42 |
| Jan-Control | 1  | 41.821  | 41.821 | 2.42 |
| Residual    | 21 | 362.977 | 17.285 |      |

| Variable | Coefficient | Standard error | T    |
|----------|-------------|----------------|------|
| Alpha    | .2178       | .878           | .25  |
| Beta     | .5993       | .385           | 1.56 |

90% Confidence Limits

|       | Lower limit | Upper limit |
|-------|-------------|-------------|
| Alpha | -1.1817     | 1.6174      |
| Beta  | -.0148      | 1.2133      |





R-squared = .2066

Standard error of estimate = 4.2804

FL

| SOURCE      | DF | AOV     |         |      |
|-------------|----|---------|---------|------|
|             |    | SS      | MS      | F    |
| Total       | 35 | 785.196 |         |      |
| Regression  | 1  | 162.252 | 162.252 | 8.86 |
| Jan-Control | 1  | 162.252 | 162.252 | 8.86 |
| Residual    | 34 | 622.945 | 18.322  |      |

| Variable | Coefficient | Standard error | T    |
|----------|-------------|----------------|------|
| Alpha    | -.2670      | .715           | -.37 |
| Beta     | 1.0095      | .339           | 2.98 |

| 90% Confidence Limits |             |             |
|-----------------------|-------------|-------------|
|                       | Lower limit | Upper limit |
| Alpha                 | -.2670      | .8543       |
| Beta                  | 1.0095      | 1.5417      |

R-squared = .0426

Standard error of estimate = 6.1508

SC

| SOURCE      | DF | AOV      |        |      |
|-------------|----|----------|--------|------|
|             |    | SS       | MS     | F    |
| Total       | 46 | 1778.131 |        |      |
| Regression  | 1  | 75.687   | 75.687 | 2.00 |
| Jan-Control | 1  | 75.687   | 75.687 | 2.00 |
| Residual    | 45 | 1702.444 | 37.832 |      |

| Variable | Coefficient | Standard error | T    |
|----------|-------------|----------------|------|
| Alpha    | -.2771      | .898           | -.31 |
| Beta     | .4355       | .308           | 1.41 |

| 90% Confidence Limits |             |             |
|-----------------------|-------------|-------------|
|                       | Lower limit | Upper limit |
| Alpha                 | -1.6766     | 1.1224      |
| Beta                  | -.0446      | .9157       |

The same model with means removed, V.19, below was tried and, although slightly better, is still not strong.

$$Y_t - \bar{Y} = \alpha + \beta(X_t - \bar{X}) + \epsilon_t \quad \text{V.19}$$



TABLE 43

ANOVA FOR REGRESSION OF SIMPLE LINEAR  
MODEL WITH MEANS REMOVED FOR ALL DATA SETS

R-squared = .1033

Standard error of estimate = 4.1575

RN

## AOV

| SOURCE      | DF | SS      | MS     | F    |
|-------------|----|---------|--------|------|
| Total       | 22 | 404.798 |        |      |
| Regression  | 1  | 41.821  | 41.821 | 2.42 |
| Jan-Control | 1  | 41.821  | 41.821 | 2.42 |
| Residual    | 21 | 362.977 | 17.285 |      |

| Variable | Coefficient | Standard error | T    |
|----------|-------------|----------------|------|
| Alpha    | 11.7215     | 1.851          | 6.33 |
| Beta     | .5993       | .385           | 1.56 |

## 90% Confidence Limits

|       | Lower limit | Upper limit |
|-------|-------------|-------------|
| Alpha | 8.7236      | 13.2483     |
|       | -.0148      | 1.2133      |

R-squared = .1747

Standard error of estimate = 4.4276

FL

## AOV

| SOURCE      | DF | SS      | MS      | F    |
|-------------|----|---------|---------|------|
| Total       | 35 | 807.618 |         |      |
| Regression  | 1  | 141.108 | 141.108 | 7.20 |
| Jan-Control | 1  | 141.108 | 141.108 | 7.20 |
| Residual    | 34 | 666.510 | 19.603  |      |

| Variable | Coefficient | Standard error | T    |
|----------|-------------|----------------|------|
| Alpha    | 10.9860     | 1.442          | 7.62 |
| Beta     | .9414       | .351           | 2.68 |

## 90% Confidence Limits

|       | Lower limit | Upper limit |
|-------|-------------|-------------|
| Alpha | 8.7236      | 13.2483     |
| Beta  | .3909       | 1.4920      |



R-squared = .5617  
 Standard error of estimate = .7589

SC

| AOV         |    |       |       |      |
|-------------|----|-------|-------|------|
| SOURCE      | DF | SS    | MS    | F    |
| Total       | 7  | 7.884 |       |      |
| Regression  | 1  | 4.428 | 4.428 | 7.69 |
| Jan-Control | 1  | 4.428 | 4.428 | 7.69 |
| Residual    | 6  | 3.456 | .576  |      |

| Variable | Coefficient | Standard error | T    |
|----------|-------------|----------------|------|
| Alpha    | -.1946      | .287           | -.68 |
| Beta     | 1.7710      | .637           | 2.77 |

| 90% Confidence Limits |             |             |
|-----------------------|-------------|-------------|
|                       | Lower limit | Upper limit |
| Alpha                 | -.7040      | .3148       |
| Beta                  | .6362       | 2.9058      |

The amount of rainfall in January does not appear to be a strong predictor for the amount of rainfall in February through December. This seems to indicate that the relationships between January rainfall and rainfall during the next eleven months is not as strong as expected. However, a further technique is available, that of log-odds and logistic regression, which are the subjects of the next section.



## VI. LOGISTIC ANALYSIS

### A. THEORY

The logistic analysis to be described in the section was developed from Gaver [Ref. 2] and Fleiss [Ref. 3]. This analysis derives from the model of I.5 as stated in the introduction.

The basic approach is to view the complement as having a binary representation, with success being defined as a complement above its mean (see equations I.3 and I.4) and failure as the complement below its mean. The problem then is to find the conditional probability of a success (the complement being above its mean for a year) given that the control (January rainfall) takes on a particular value.

The control is now taken to be the logged rainfall anomaly of January and is found in equation I.2 repeated below.

$$X_t = \ln(R_{t,4}) - \frac{1}{N} \sum_{t=1}^N \ln(R_{t,4}) \quad \text{VI.1}$$

If the probability of success, given  $X_t$  is written as

$$P(\text{Success} | X_t) = \theta_t \quad \text{VI.2}$$

a superficially attractive model for  $\theta_t$  is

$$\theta_t = \alpha + \beta X_t + \varepsilon_t \quad \text{VI.3}$$

This model has two difficulties, the worst of which is that probabilities of greater than one or less than zero are





allowed. Secondly, the  $\theta_t$  are not available in proportion form to fit the model.

Initially, the problem of estimating  $\theta_t$  is approached by grouping the data. Section II indicated that each year seemed to be independent of the next and that there was no trend. This allows for the ordering of the  $X_t$ 's into their order statistics  $X_{(t)}$ 's. Once the ordering has been done, non overlapping groups of arbitrary size may be formed as shown. Let  $X_1, X_2, X_3, \dots, X_{12}$  be a series of 12 years from an arbitrary data set, with associated order statistics  $X_{(1)}, X_{(2)}, X_{(3)}, \dots, X_{(12)}$ . If groups of size three are desired, the data are partitioned below.

$$X_{(1)}, X_{(2)}, X_{(3)} | X_{(4)}, X_{(5)}, X_{(6)} | \dots | X_{(10)}, X_{(11)}, X_{(12)}$$

Given these groups, let  $\tilde{X}_j$  be a measure of location for the  $j^{\text{th}}$  group. This analysis used the median, therefore  $\tilde{X}_j = X_{(3j+1)}$ . Also associated with each group is  $R_j$  (not rainfall), the number of success in group  $j$ , and  $n_j$ , the number of elements in group  $j$ . From this set up, the required probabilities may be estimated as;

$$\hat{\theta}_j = R_j/n_j \quad \text{VI.4}$$

A solution to the first problem, that of the model yielding probabilities outside of  $(0,1)$ , is to use the log odds, instead of  $\theta_j$  where;

$$\text{Log odds} = \phi_j = \ln \left( \frac{\theta_j}{1-\theta_j} \right) \quad \text{VI.5}$$



which is equivalent to the logarithm of the odds ratio as given in V.13. Gaver [Ref. 2] suggests that a correction of .5 be applied to guard against the problem of 0 and 1 within the logarithm and to reduce the bias. The statistic then becomes;

$$\phi_j' = \ln \left( \frac{\theta_j + .5}{1.5 - \theta_j} \right) \quad \text{VI.6}$$

The temptation is to go directly to the model

$$\phi_j' = \alpha + \beta \tilde{X}_j; \quad \text{VI.7}$$

yet  $V[\theta_j] = \theta_j(1-\theta_j)/n_j$  which is not constant. This suggests the need for a weighting scheme.

The weighting scheme used was that of iteratively reweighted least squares, (IRWLS), using the bi-weights. This method is discussed in detail in Mosteller and Tukey [Ref. 13].

Although grouping of the data and the model of VI.7 provide adequate representation of the underlying structure, the logistic model itself, I.5, when viewed through the eyes of maximum likelihood theory can yield more insight.

The model is assumed to be

$$\theta_t = \frac{e^{\alpha + \beta X_t}}{1 + e^{\alpha + \beta X_t}} \quad \text{VI.8}$$

where the  $X_t$  are independent.



The likelihood function is then

$$\begin{aligned}
 L(X, Y; \alpha, \beta) &= \prod_{t=1}^N \left( \frac{e^{\alpha + \beta X_t}}{1 + e^{\alpha + \beta X_t}} \right)^{Y_t} \left( \frac{1}{1 + e^{\alpha + \beta X_t}} \right)^{1 - Y_t} \\
 &= \prod_{t=1}^N \frac{e^{\alpha Y_t + \beta X_t Y_t}}{\prod_{t=1}^N (1 + e^{\alpha + \beta X_t})}
 \end{aligned} \tag{V.9}$$

and the log-likelihood is

$$\begin{aligned}
 L(X, Y; \alpha, \beta) &= \alpha \sum_{t=1}^N Y_t + \beta \sum_{t=1}^N X_t Y_t \\
 &\quad - \sum_{t=1}^N \ln (1 + e^{\alpha + \beta X_t})
 \end{aligned} \tag{V.10}$$

The gradient, and Hessian of VI.9 are

$$\frac{\partial L}{\partial \alpha} = \sum_{t=1}^N Y_t - \sum_{t=1}^N \left( \frac{\psi_t}{1 + \psi_t} \right)$$

$$\frac{\partial L}{\partial \beta} = \sum_{t=1}^N X_t Y_t - \sum_{t=1}^N X_t \left( \frac{\psi_t}{1 + \psi_t} \right)$$

and

$$H_L = \begin{pmatrix} - \sum_{t=1}^N \frac{\psi_t}{(1 + \psi_t)^2} & - \sum_{t=1}^N \frac{X_t \psi_t}{(1 + \psi_t)^2} \\ - \sum_{t=1}^N \frac{X_t \psi_t}{(1 + \psi_t)^2} & - \sum_{t=1}^N \frac{X_t^2 \psi_t}{(1 + \psi_t)^2} \end{pmatrix} \tag{V.11}$$

where

$$\psi_t = e^{\alpha + \beta X_t}.$$



A simple way to solve for  $\hat{\alpha}$  and  $\hat{\beta}$  is to use Newtons Method as;

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix}_{k+1} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}_k - H_L^{-1} \nabla t_L$$

since all necessary elements may be calculated in one pass of the computer algorithm.

One beneficial byproduct of the maximum likelihood approach is the asymptotic information matrix,  $H_L^{-1}$ . Gaver [Ref. 2] states that the diagonal elements of this matrix provide good estimates of  $V[\hat{\alpha}]$  and  $V[\hat{\beta}]$  under assumptions of normality.

## B. ANALYSIS

### 1. Grouped Data

The first approach taken was that of grouping the data as described above. Groups of 3, 4, and 5 were used, as were two separate methods of regression, ordinary least squares (OLS), and iteratively reweighted least squares (IRWLS). Tables 44 (RN), 45 (FL), and 46 (SC) present the data and Tables 47 (OLS), and 48 (IRWLS) present the results of the regressions.





TABLE 44

DATA SET RN, LOGGED JANUARY ANOMALIES AND  
SUCCESES FOR GROUPED AND UNGROUPED FORMS

| UNGROUPED |       |   | GROUPED |         |   |
|-----------|-------|---|---------|---------|---|
| t         | X     | Y | j       | X       | Y |
| 1         | -1.22 | 0 |         | group=3 |   |
| 2         | -1.09 | 0 | 1       | -1.09   | 0 |
| 3         | - .66 | 0 | 2       | - .46   | 1 |
| 4         | - .58 | 0 | 3       | - .19   | 1 |
| 5         | - .46 | 0 | 4       | .01     | 2 |
| 6         | - .36 | 1 | 5       | .16     | 2 |
| 7         | - .31 | 0 | 6       | .36     | 2 |
| 8         | - .18 | 1 | 7       | .50     | 2 |
| 9         | - .17 | 0 | 8       | .98     | 2 |
| 10        | - .05 | 1 |         | group=4 |   |
| 11        | .01   | 1 | 1       | -.38    | 0 |
| 12        | .01   | 0 | 2       | -.34    | 2 |
| 13        | .15   | 1 | 3       | -.03    | 2 |
| 14        | .16   | 0 | 4       | .20     | 2 |
| 15        | .24   | 1 | 5       | .47     | 4 |
| 16        | .28   | 0 | 6       | .94     | 2 |
| 17        | .36   | 1 |         | group=5 |   |
| 18        | .46   | 1 | 1       | -.66    | 0 |
| 19        | .48   | 1 | 2       | -.19    | 3 |
| 20        | .50   | 1 | 3       | .15     | 3 |
| 21        | .51   | 0 | 4       | .46     | 4 |
| 22        | .94   | 1 | 5       | .94     | 2 |
| 23        | 1.01  |   |         |         |   |



TABLE 45

DATA SET FL, LOGGED JANUARY ANOMALIES AND  
SUCCESES FOR GROUPED AND UNGROUPED DATA

| UNGROUPED |       |   | GROUPED |         |   |
|-----------|-------|---|---------|---------|---|
| t         | X     | Y | j       | X       | Y |
| 1         | -3.41 | 1 |         | group=3 |   |
| 2         | -1.82 | 0 | 1       | -1.82   | 1 |
| 3         | -1.14 | 0 | 2       | -1.00   | 0 |
| 4         | -1.02 | 0 | 3       | - .46   | 0 |
| 5         | -1.00 | 0 | 4       | - .20   | 2 |
| 6         | - .92 | 0 | 5       | .11     | 1 |
| 7         | - .47 | 0 | 6       | .14     | 1 |
| 8         | - .46 | 0 | 7       | .20     | 0 |
| 9         | - .39 | 0 | 8       | .34     | 2 |
| 10        | - .28 | 1 | 9       | .50     | 2 |
| 11        | - .20 | 0 | 10      | .61     | 2 |
| 12        | - .14 | 1 | 11      | .75     | 3 |
| 13        | - .04 | 1 | 12      | 1.14    | 3 |
| 14        | .11   | 0 |         | group=4 |   |
| 15        | .11   | 0 | 1       | -1.48   | 1 |
| 16        | .13   | 0 | 2       | - .70   | 0 |
| 17        | .14   | 0 | 3       | - .24   | 2 |
| 18        | .16   | 1 | 4       | .11     | 1 |
| 19        | .16   | 0 | 5       | .16     | 1 |
| 20        | .20   | 0 | 6       | .32     | 2 |
| 21        | .25   | 0 | 7       | .50     | 2 |
| 22        | .30   | 1 | 8       | .71     | 4 |
| 23        | .34   | 1 | 9       | 1.11    | 4 |
| 24        | .38   | 0 |         | group=5 |   |
| 25        | .46   | 1 | 1       | -1.14   | 1 |
| 26        | .50   | 0 | 2       | - .46   | 1 |
| 27        | .51   | 1 | 3       | - .04   | 2 |
| 28        | .57   | 0 | 4       | .16     | 1 |
| 29        | .61   | 1 | 5       | .34     | 3 |
| 30        | .71   | 1 | 6       | .57     | 3 |
| 31        | .71   | 1 | 7       | .95     | 6 |
| 32        | .75   | 1 |         |         |   |
| 33        | .82   | 1 |         |         |   |
| 34        | 1.08  | 1 |         |         |   |
| 35        | 1.14  | 1 |         |         |   |
| 36        | 1.31  | 1 |         |         |   |



TABLE 46

DATA SET SC, LOGGED JANUARY ANOMALIES AND  
SUCCESSIONS FOR GROUPED AND UNGROUPED FORMS

| UNGROUPED |       |   | GROUPED |         |   |
|-----------|-------|---|---------|---------|---|
| t         | X     | Y | j       | X       | Y |
| 1         | -4.24 | 1 |         | group=3 |   |
| 2         | -1.41 | 0 | 1       | -1.41   | 1 |
| 3         | -1.38 | 0 | 2       | -1.07   | 0 |
| 4         | -1.08 | 0 | 3       | - .96   | 1 |
| 5         | -1.07 | 0 | 4       | - .55   | 1 |
| 6         | -1.06 | 0 | 5       | - .26   | 2 |
| 7         | -1.02 | 0 | 6       | - .10   | 1 |
| 8         | - .96 | 1 | 7       | .02     | 2 |
| 9         | - .65 | 0 | 8       | .20     | 0 |
| 10        | - .60 | 0 | 9       | .28     | 2 |
| 11        | - .55 | 1 | 10      | .40     | 2 |
| 12        | - .36 | 0 | 11      | .46     | 1 |
| 13        | - .27 | 1 | 12      | .56     | 2 |
| 14        | - .26 | 0 | 13      | .69     | 0 |
| 15        | - .21 | 1 | 14      | .78     | 2 |
| 16        | - .17 | 1 | 15      | .97     | 1 |
| 17        | - .10 | 0 | 16      | 1.25    | 2 |
| 18        | - .03 | 0 |         | group=4 |   |
| 19        | - .01 | 1 | 1       | -1.39   | 1 |
| 20        | .02   | 0 | 2       | -1.04   | 1 |
| 21        | .12   | 1 | 3       | - .58   | 1 |
| 22        | .14   | 0 | 4       | .23     | 3 |
| 23        | .20   | 0 | 5       | - .02   | 1 |
| 24        | .22   | 0 | 6       | .17     | 1 |
| 25        | .23   | 0 | 7       | .28     | 3 |
| 26        | .28   | 1 | 8       | .42     | 2 |
| 27        | .28   | 1 | 9       | .56     | 2 |
| 28        | .40   | 1 | 10      | .69     | 1 |
| 29        | .40   | 1 | 11      | .78     | 1 |
| 30        | .41   | 0 | 12      | 1.03    | 3 |
| 31        | .43   | 0 |         | group=5 |   |
| 32        | .46   | 1 | 1       | -1.38   | 1 |
| 33        | .49   | 0 | 2       | - .96   | 1 |
| 34        | .56   | 1 | 3       | - .27   | 3 |
| 35        | .56   | 1 | 4       | - .04   | 2 |
| 36        | .60   | 0 | 5       | .20     | 1 |
| 37        | .64   | 0 | 6       | .40     | 4 |
| 38        | .69   | 0 | 7       | .49     | 3 |
| 39        | .70   | 0 | 8       | .69     | 1 |
| 40        | .70   | 1 | 9       | .88     | 2 |
| 41        | .78   | 1 | 10      | 1.25    | 2 |
| 42        | .81   | 0 |         |         |   |
| 43        | .88   | 0 |         |         |   |
| 44        | .97   | 0 |         |         |   |
| 45        | .97   | 1 |         |         |   |
| 46        | 1.03  | 1 |         |         |   |
| 47        | 1.47  | 1 |         |         |   |



TABLE 47a

ORDINARY LEAST SQUARES REGRESSION WITH  
THE MODEL OF EQUATION VI.7 FOR DATA SET RN

R-squared = .0423

Standard error of estimate = 6.1508

RN  
GROUP=3

| SOURCE      | DF | SS       | MS     | F    |
|-------------|----|----------|--------|------|
| Total       | 46 | 1778.131 |        |      |
| Regression  | 1  | 75.687   | 75.687 | 2.00 |
| Jan-Control | 1  | 75.687   | 75.687 | 2.00 |
| Residual    | 45 | 1702.444 | 37.832 |      |

| Variable | Coefficient | Standard error | T    |
|----------|-------------|----------------|------|
| Alpha    | 14.5833     | 1.677          | 8.69 |
| Beta     | .4355       | .308           | 1.41 |

## 90% Confidence Limits

|       | Lower limit | Upper limit |
|-------|-------------|-------------|
| Alpha | 11.9680     | 17.1988     |
| Beta  | -.0446      | .9157       |

R-squared = .4068

Standard error of estimate = 1.2099

RN  
GROUP=4

| SOURCE      | DF | SS    | MS    | F    |
|-------------|----|-------|-------|------|
| Total       | 5  | 9.871 |       |      |
| Regression  | 1  | 4.016 | 4.016 | 2.74 |
| Jan-Control | 1  | 4.016 | 4.016 | 2.74 |
| Residual    | 4  | 5.855 | 1.464 |      |

| Variable | Coefficient | Standard error | T    |
|----------|-------------|----------------|------|
| Alpha    | -.1696      | .517           | -.33 |
| Beta     | 1.7687      | 1.068          | 1.66 |

## 90% Confidence Limits

|       | Lower limit | Upper limit |
|-------|-------------|-------------|
| Alpha | -1.1675     | .8284       |
| Beta  | 1.7688      | 3.8288      |





R-squared = .5973  
 Standard error of estimate = .9995

RN  
 GROUP=5

| AOV                   |             |       |                |      |
|-----------------------|-------------|-------|----------------|------|
| SOURCE                | DF          | SS    | MS             | F    |
| Total                 | 4           | 7.442 |                |      |
| Regression            | 1           | 4.446 | 4.446          | 4.45 |
| Jan-Control           | 1           | 4.446 | 4.446          | 4.45 |
| Residual              | 3           | 2.997 | .999           |      |
| Variable              | Coefficient |       | Standard error | T    |
| Alpha                 | -.2647      |       | .461           | -.57 |
| Beta                  | 1.7193      |       | .815           | 2.11 |
| 90% Confidence Limits |             |       |                |      |
|                       | Lower limit |       | Upper limit    |      |
| Alpha                 | -1.2358     |       | .7064          |      |
| Beta                  | .0040       |       | 3.4345         |      |

TABLE 47b

ORDINARY LEAST SQUARES REGRESSION WITH  
 THE MODEL OF EQUATION VI.7 FOR DATA SET FL

R-squared = .2862  
 Standard error of estimate = .8949

FL  
 GROUP=3

| AOV                   |             |        |                |      |
|-----------------------|-------------|--------|----------------|------|
| SOURCE                | DF          | SS     | MS             | F    |
| Total                 | 11          | 20.591 |                |      |
| Regression            | 1           | 7.998  | 7.998          | 6.35 |
| Jan-Control           | 1           | 7.998  | 7.998          | 6.35 |
| Residual              | 10          | 12.593 | 1.259          |      |
| Variable              | Coefficient |        | Standard error | T    |
| Alpha                 | -.1459      |        | .324           | -.33 |
| Beta                  | 1.0521      |        | .417           | 2.52 |
| 90% Confidence Limits |             |        |                |      |
|                       | Lower limit |        | Upper limit    |      |
| Alpha                 | -.6871      |        | .4944          |      |
| Beta                  | .3550       |        | 1.7492         |      |



R-squared = .5370  
 Standard error of estimate = 1.0484

FL  
 GROUP=4

# AOV

| SOURCE      | DF | SS     | MS    | F    |
|-------------|----|--------|-------|------|
| Total       | 8  | 16.615 |       |      |
| Regression  | 1  | 8.922  | 8.922 | 8.12 |
| Jan-Control | 1  | 8.922  | 8.922 | 8.12 |
| Residual    | 7  | 7.963  | 1.099 |      |

| Variable | Coefficient | Standard Error | T    |
|----------|-------------|----------------|------|
| Alpha    | -.1148      | .350           | -.33 |
| Beta     | 1.3572      | .476           | 2.83 |

## 90% Confidence Limits

|       | Lower limit | Upper limit |
|-------|-------------|-------------|
| Alpha | -.7237      | .4941       |
| Beta  | .5297       | 2.1847      |

R-squared = .6311  
 Standard error of estimate = .8809

FL  
 GROUP=5

# AOV

| SOURCE      | DF | SS     | MS    | F    |
|-------------|----|--------|-------|------|
| Total       | 6  | 10.519 |       |      |
| Regression  | 1  | 6.639  | 6.639 | 8.56 |
| Jan-Control | 1  | 6.639  | 6.639 | 8.56 |
| Residual    | 5  | 3.880  | .776  |      |

| Variable | Coefficient | Standard Error | T    |
|----------|-------------|----------------|------|
| Alpha    | -.1393      | .334           | -.42 |
| Beta     | 1.5225      | .521           | 2.92 |

## 90% Confidence Limits

|       | Lower limit | Upper limit |
|-------|-------------|-------------|
| Alpha | -.7525      | .4740       |
| Beta  | .5673       | 2.4777      |



TABLE 47c

ORDINARY LEAST SQUARES REGRESSION WITH  
THE MODEL OF EQUATION VI.7 FOR DATA SET SC

R-squared = .1404

Standard error of estimate - .9901

SC

GROUP=3

## AOV

| SOURCE      | DF | SS     | MS    | F    |
|-------------|----|--------|-------|------|
| Total       | 15 | 15.965 |       |      |
| Regression  | 1  | 2.241  | 2.241 | 2.29 |
| Jan-Control | 1  | 2.241  | 2.241 | 2.29 |
| Residual    | 14 | 13.724 | .980  |      |

| Variable | Coefficient | Standard Error | T     |
|----------|-------------|----------------|-------|
| Alpha    | -.3039      | .249           | -1.22 |
| Beta     | .5059       | .335           | 1.51  |

## 90% Confidence Limits

|       | Lower limit | Upper limit |
|-------|-------------|-------------|
| Alpha | -.7090      | .1012       |
| Beta  | -.0386      | 1.0504      |

R-squared = .1872

Standard error of estimate = .8959

SC

GROUP=4

## AOV

| SOURCE      | DF | SS    | MS    | F    |
|-------------|----|-------|-------|------|
| Total       | 11 | 9.875 |       |      |
| Regression  | 1  | 1.848 | 1.848 | 2.30 |
| Jan-Control | 1  | 1.848 | 1.848 | 2.30 |
| Residual    | 10 | 8.026 | .803  |      |

| Variable | Coefficient | Standard Error | T    |
|----------|-------------|----------------|------|
| Alpha    | -.2284      | .260           | -.88 |
| Beta     | .5447       | .359           | 1.52 |

## 90% Confidence Limits

|       | Lower limit | Upper limit |
|-------|-------------|-------------|
| Alpha | -.6618      | .2051       |
| Beta  | .5447       | 1.1439      |



R=squared = .2862  
 Standard error of estimate = .8949

SC  
 GROUP=5

| ADV         |    |       |       |      |
|-------------|----|-------|-------|------|
| SOURCE      | DF | SS    | MS    | F    |
| Total       | 9  | 8.976 |       |      |
| Regression  | 1  | 2.569 | 2.569 | 3.21 |
| Jan-Control | 1  | 2.569 | 2.569 | 3.21 |
| Residual    | 8  | 6.407 | .801  |      |

| Variable | Coefficient | Standard Error | T    |
|----------|-------------|----------------|------|
| Alpha    | -.1838      | .287           | -.64 |
| Beta     | .6573       | .367           | 1.79 |

| 90% Confidence Limits |             |             |
|-----------------------|-------------|-------------|
|                       | Lower limit | Upper limit |
| Alpha                 | -.6736      | .3061       |
| Beta                  | .0303       | 1.2843      |

TABLE 48

ITERATIVELY REWEIGHTED LEAST SQUARES  
 REGRESSION USING BI-WEIGHTS FOR THE MODEL OF EQUATION VI.7

| GROUP SIZE 3 |    |                |               |
|--------------|----|----------------|---------------|
| C=9          |    | $\hat{\alpha}$ | $\hat{\beta}$ |
| Data sets    | RN | .048           | 1.013         |
|              | FL | -.063          | .664          |
|              | SC | -.116          | .351          |

|           |    |       |       |
|-----------|----|-------|-------|
| C=4       |    |       |       |
| Data sets | RN | .052  | 1.031 |
|           | FL | -.197 | 1.049 |
|           | SC | .154  | .480  |

| GROUP SIZE 4 |    |       |      |
|--------------|----|-------|------|
| C=9          |    |       |      |
| Data sets    | RN | -.103 | .665 |
|              | FL | -.093 | .705 |
|              | SC | -.169 | .280 |

| GROUP SIZE 5 |    |       |      |
|--------------|----|-------|------|
| C=9          |    |       |      |
| Data sets    | RN | -.065 | .865 |
|              | FL | -.107 | .723 |
|              | SC | -.126 | .413 |





## 2. Maximum Likelihood

Table 49 displays the final points for each data set along with the inverse Hessian at that point. Figures 103 (RN), 104 (FL), and 105 (SC) are interesting in that they portray the contours of the likelihood functions for each data set. These contours show the variance of the estimated parameters in a graphic way. Note how data sets RN and FL seem to have some sort of horizontal ridge indicating a good pick of the slope parameter, yet the contours of data set SC are almost circular about the origin of the axes indicating no significant difference from zero for either parameter.

TABLE 49

MAXIMUM LIKELIHOOD ESTIMATES OF  $\hat{\alpha}$  AND  $\hat{\beta}$  ALONG WITH ESTIMATES OF THEIR VARIANCE FOR ALL THREE DATA SETS

|                        |                          |  |
|------------------------|--------------------------|--|
| Data set RN            |                          |  |
| $\hat{\alpha} = .062$  |                          | $H_L^{-1} = \begin{pmatrix} .257 & -.043 \\ -.034 & 1.720 \end{pmatrix}$ |
| $\hat{\beta} = 2.918$  |                          |  |
|                        | $V[\hat{\alpha}] = .257$ |  |
|                        | $V[\hat{\beta}] = 1.720$ |  |
| Data set FL            |                          |  |
| $\hat{\alpha} = -.171$ |                          | $H_L^{-1} = \begin{pmatrix} .129 & -.040 \\ -.040 & .298 \end{pmatrix}$  |
| $\hat{\beta} = .933$   |                          |  |
|                        | $V[\hat{\alpha}] = .129$ |  |
|                        | $V[\hat{\beta}] = .298$  |  |
| Data set SC            |                          |  |
| $\hat{\alpha} = -.303$ |                          | $H_L^{-1} = \begin{pmatrix} .088 & -.035 \\ -.035 & .113 \end{pmatrix}$  |
| $\hat{\beta} = .171$   |                          |  |
|                        | $V[\hat{\alpha}] = .088$ |  |
|                        | $V[\hat{\beta}] = .113$  |  |



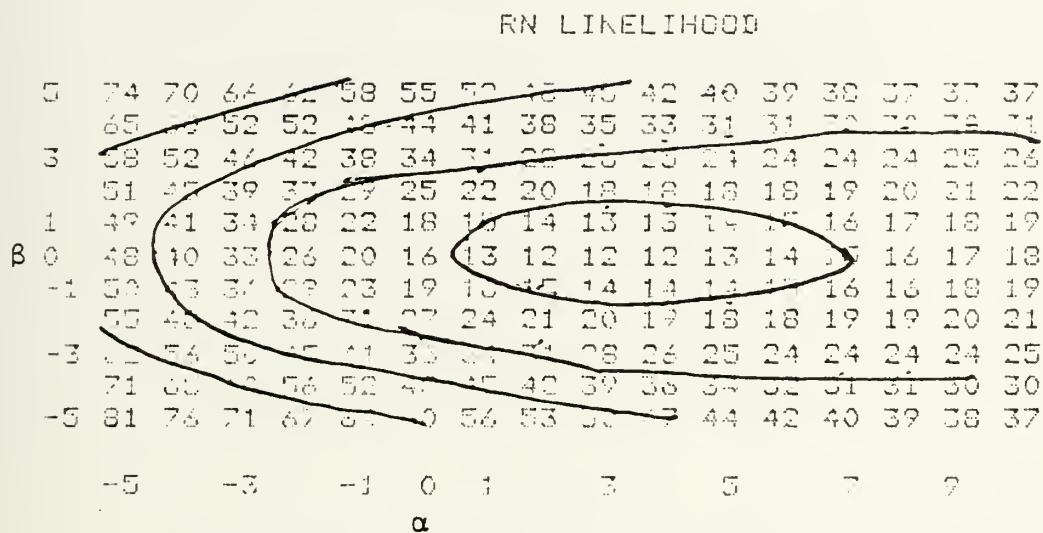


Figure 103. Contours of log likelihood function for data set RN

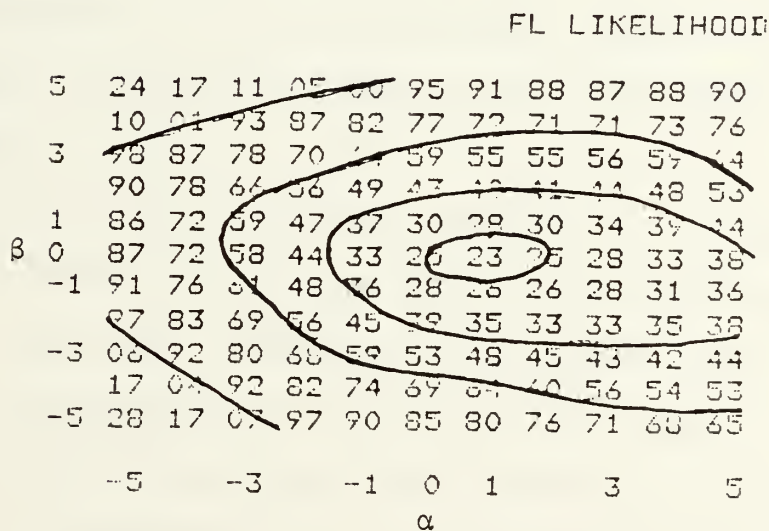


Figure 104. Contours of log likelihood function for data set FL



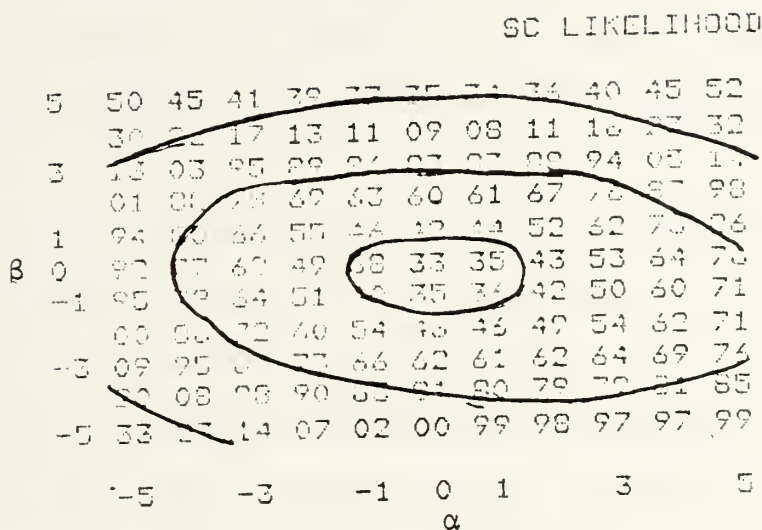


Figure 105. Contours of log likelihood function for data set SC

### C. DISCUSSION

A recapitulation of all the parameter values is in Table 50. Some interesting observations to be made from this table are:

(1) The slope for data set RN, as found by maximum likelihood, is much greater than that of any other method or data set. The first temptation is to treat this as an outlier, yet the evidence of the contour plot and of the validation of the next section tend to back up this number. The reason for this difference is a possible subject for further reserach.

(2) Except for the intercept values, the slopes of data sets RN and FL seem to be fairly consistent within and between



regression methods. This comment is made in light of the difference of these data sets and that of SC.

(3) Data sets RN and FL seem to be similar in many ways, yet data set SC appears to be different in both degree and significance.

No other strong pattern is apparent in these parameter values. Graphical displays of the parameters, as used with grouped data are given in Figures 106 (RN), 107 (FL), and 108 (SC). Again, note the significant difference of the maximum likelihood line for data set RN. Table 51 contains the data points from which Figures 106, 107 and 108 were drawn.

After viewing these figures, the maximum likelihood approach is the preferred method for the Peninsula data sets, whereas the robust( $C=4$ ) IRWLS may be best for the Valley data set. Although fits were made to data set SC, it appears as if no great significance has been found.





TABLE 50  
PARAMETER FIT RECAPITULATION  
FOR ALL DATA SETS

|        |                    | DATA SETS  |       |       |
|--------|--------------------|------------|-------|-------|
| Method |                    | RN         | FL    | SC    |
| 1.     | Maximum Likelihood | $\alpha$ : | .062  | -.171 |
|        |                    | $\beta$ :  | 2.918 | .933  |
|        |                    |            |       | -.303 |
|        |                    |            |       | .171  |
| 2.     | OLS                |            |       |       |
|        | Group=3            | $\alpha$ : | -.195 | -.146 |
|        |                    | $\beta$ :  | 1.771 | 1.052 |
|        |                    |            |       | -.304 |
|        |                    |            |       | .506  |
|        | Group=4            | $\alpha$ : | -.170 | -.115 |
|        |                    | $\beta$ :  | 1.769 | 1.357 |
|        |                    |            |       | -.228 |
|        |                    |            |       | .545  |
|        | Group=5            | $\alpha$ : | -.265 | -.139 |
|        |                    | $\beta$ :  | 1.719 | 1.523 |
|        |                    |            |       | -.184 |
|        |                    |            |       | .657  |
| 3.     | IRWLS              |            |       |       |
|        | Group=3            | $\alpha$ : | .048  | -.063 |
|        | C=9                | $\beta$ :  | 1.013 | .664  |
|        |                    |            |       | .116  |
|        |                    |            |       | .351  |
|        | Group=3            | $\alpha$ : | .052  | -.197 |
|        | C=4                | $\beta$ :  | 1.031 | 1.049 |
|        |                    |            |       | .154  |
|        |                    |            |       | .480  |
|        | Group 4            | $\alpha$ : | -.103 | -.093 |
|        | C=9                | $\beta$ :  | .665  | .705  |
|        |                    |            |       | .169  |
|        |                    |            |       | .280  |
|        | Group=5            | $\alpha$ : | -.065 | -.107 |
|        | C=9                | $\beta$ :  | .855  | .723  |
|        |                    |            |       | .126  |
|        |                    |            |       | .413  |



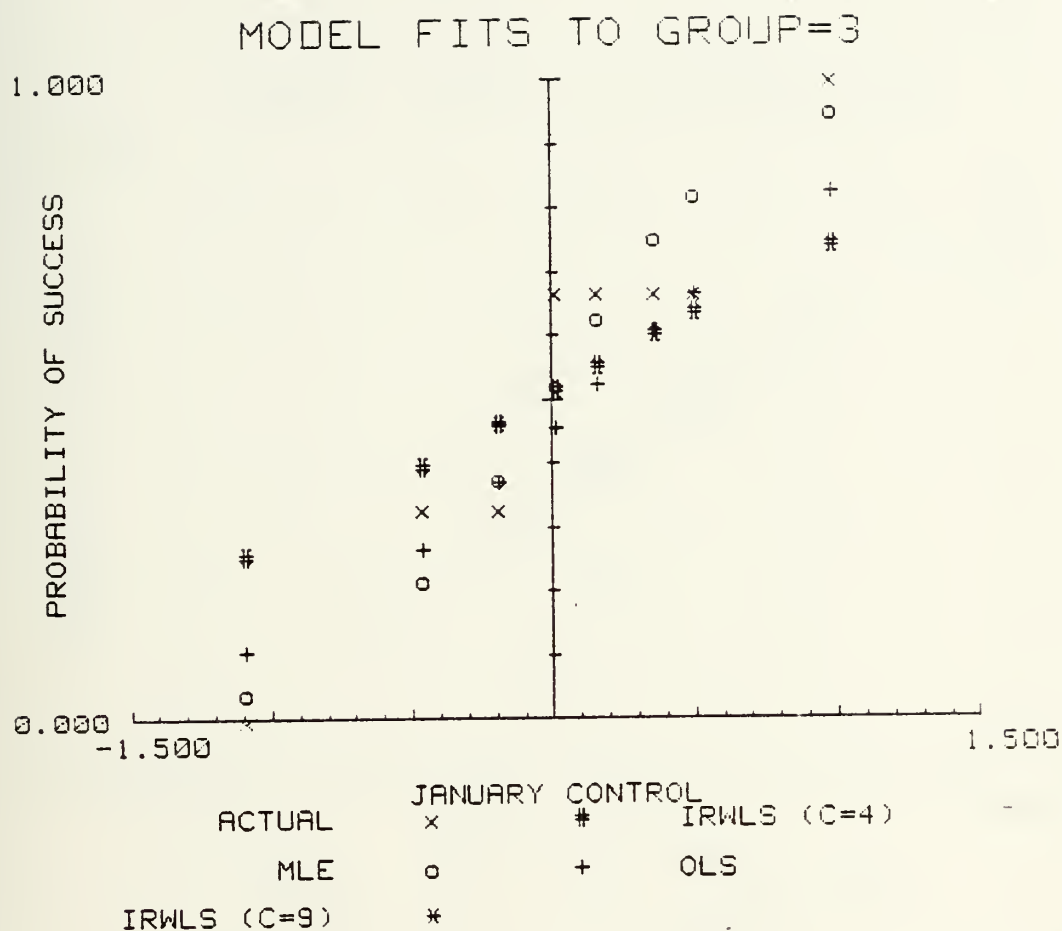


Figure 106. Estimated probability of greater-than-average total rest-of-year rainfall versus the anomaly of logged rainfall for January for data set RN



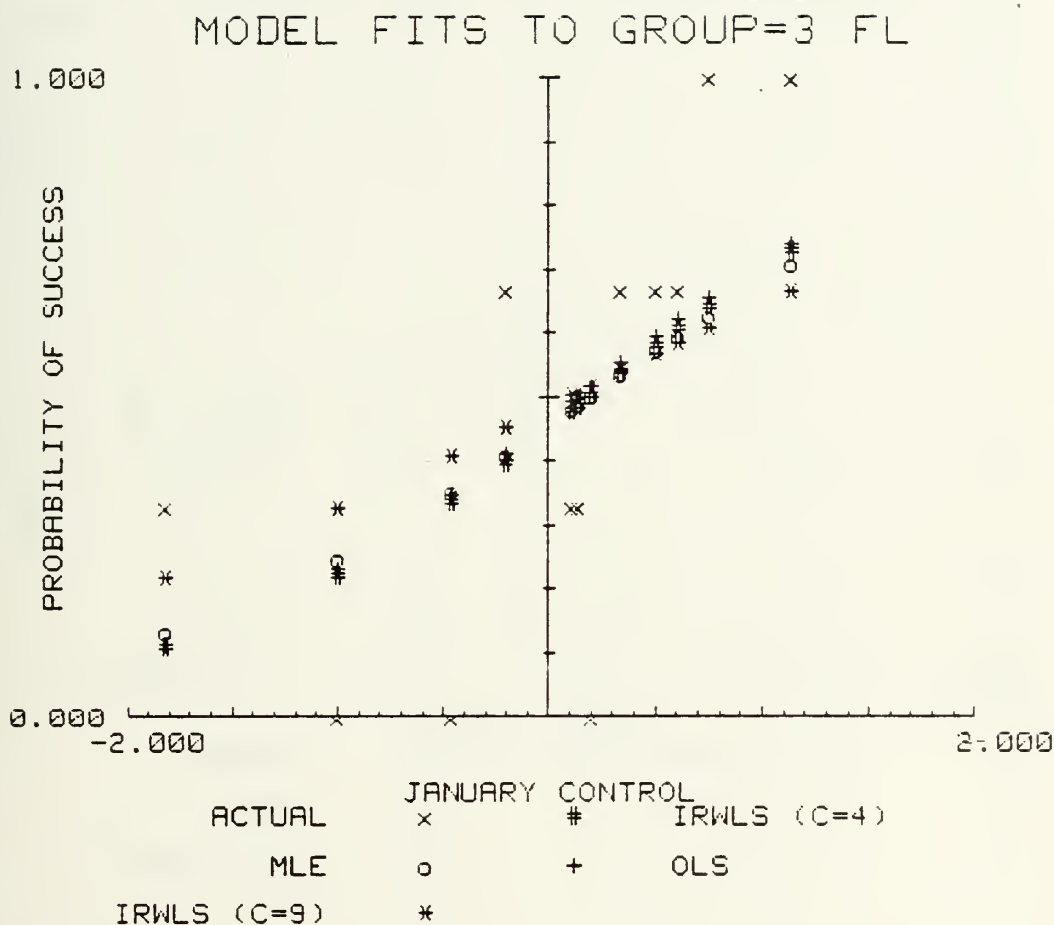


Figure 107. Estimated probability of greater-than-average total rest-of-year rainfall versus the anomaly of logged rainfall for January for data set FL



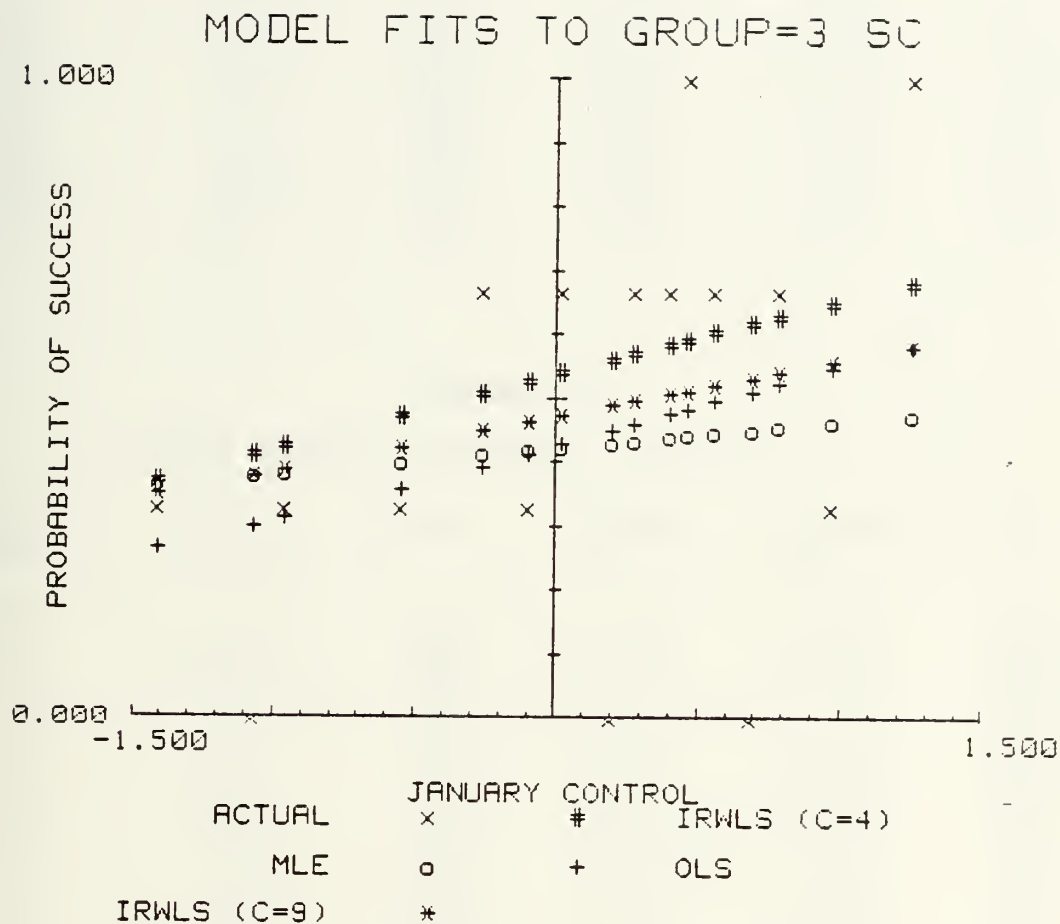


Figure 108. Estimated probability of greater-than average total rest-of-year rainfall versus the anomaly of logged rainfall for January for data set FL





TABLE 51a

## ACTUAL VALUES FOR MODEL FITS OF RN

| LOGGED<br>ANOMALY | ACTUAL<br>VALUE | MLE  | IRWLS<br>C=9 | IRWLS<br>C=4 | OLS  |
|-------------------|-----------------|------|--------------|--------------|------|
| -1.090            | 0.000           | .042 | .258         | .255         | .107 |
| - .460            | .330            | .218 | .397         | .396         | .267 |
| - .190            | .330            | .379 | .464         | .464         | .370 |
| .010              | .670            | .523 | .515         | .516         | .456 |
| .160              | .670            | .629 | .552         | .554         | .522 |
| .360              | .670            | .753 | .602         | .604         | .609 |
| .500              | .670            | .821 | .635         | .638         | .666 |
| .980              | 1.000           | .949 | .739         | .743         | .824 |

TABLE 51b

## ACTUAL VALUES FOR MODEL FITS OF FL

| LOGGED<br>ANOMALY | ACTUAL<br>VALUE | MLE  | IRWLS<br>C=9 | IRWLS<br>C=4 | OLS  |
|-------------------|-----------------|------|--------------|--------------|------|
| -1.820            | .330            | .124 | .219         | .108         | .113 |
| -1.000            | 0.000           | .249 | .326         | .223         | .232 |
| - .460            | 0.000           | .354 | .409         | .336         | .348 |
| - .200            | .670            | .412 | .451         | .400         | .412 |
| .110              | .330            | .483 | .503         | .480         | .492 |
| .140              | .330            | .490 | .507         | .487         | .500 |
| .200              | 0.000           | .504 | .517         | .503         | .516 |
| .340              | .670            | .536 | .541         | .540         | .553 |
| .500              | .670            | .573 | .567         | .581         | .594 |
| .610              | .670            | .598 | .585         | .609         | .621 |
| .750              | 1.000           | .629 | .607         | .643         | .655 |
| 1.140             | 1.000           | .709 | .667         | .731         | .741 |



TABLE 51c

## ACTUAL VALUES FOR MODEL FITS OF SC

| LOGGED<br>ANOMALY | ACTUAL<br>VALUE | MEL  | IRWLS<br>C=9 | IRWLS<br>C=4 | OLS  |
|-------------------|-----------------|------|--------------|--------------|------|
| -1.410            | .330            | .367 | .352         | .372         | .266 |
| -1.070            | 0.000           | .381 | .380         | .411         | .300 |
| - .960            | .330            | .385 | .389         | .424         | .312 |
| - .550            | .330            | .402 | .423         | .473         | .358 |
| - .260            | .670            | .414 | .448         | .507         | .393 |
| - .100            | .330            | .421 | .462         | .526         | .412 |
| .020              | .670            | .426 | .473         | .541         | .427 |
| .200              | 0.000           | .433 | .489         | .562         | .449 |
| .280              | .670            | .437 | .496         | .572         | .460 |
| .400              | .670            | .442 | .506         | .586         | .475 |
| .460              | 1.000           | .444 | .511         | .593         | .482 |
| .560              | .670            | .448 | .520         | .604         | .495 |
| .690              | 0.000           | .454 | .532         | .619         | .511 |
| .780              | .670            | .458 | .539         | .629         | .523 |
| .970              | .330            | .466 | .556         | .650         | .547 |
| 1.250             | 1.000           | .478 | .580         | .680         | .581 |



## VII. VALIDATION OF LOGISTIC MODELS

### A. GENERAL

The various parameters that were estimated in the previous section all may be subject to some sort of validation. However, this paper will only view the validation for the maximum likelihood approach on all data sets and the IRWLS (C=4) approach on data set SC. The validation will be conducted against the reserved, independent, data sets of years 1975 through 1980. These are the same data sets as used in section IV.

Table 52 portrays the reserved data, in a form for logistic analysis, and Figure 109 is a display of the derived contingency tables for the reserved data.



TABLE 52

RESERVED DATA IN FORM FOR THE  
LOGISTIC ANALYSIS

| YEAR | DATA SET RN<br>X | COMPLEMENT | Y <sub>t</sub> |
|------|------------------|------------|----------------|
| 1975 | -.739            | 12.92      | 0              |
| 1976 | -2.446           | 11.28      | 0              |
| 1977 | -.477            | 11.14      | 0              |
| 1978 | .883             | 19.62      | 1              |
| 1979 | .542             | 18.11      | 1              |
| 1980 | .752             | ---        | -              |

| YEAR | DATA SET FL<br>X | COMPLEMENT | Y <sub>t</sub> |
|------|------------------|------------|----------------|
| 1975 | -.470            | 12.58      | 0              |
| 1976 | -2.610           | 10.58      | 0              |
| 1977 | -.470            | 11.10      | 0              |
| 1978 | .734             | 19.32      | 1              |
| 1979 | .547             | 13.72      | 0              |
| 1980 | .633             | ---        | -              |

| YEAR | DATA SET SC<br>X | COMPLEMENT | Y <sub>t</sub> |
|------|------------------|------------|----------------|
| 1975 | -.535            | 17.97      | 1              |
| 1976 | -3.773           | 11.19      | 0              |
| 1977 | .352             | 10.03      | 0              |
| 1978 | 1.054            | 23.38      | 1              |
| 1979 | .512             | 16.36      | 0              |
| 1980 | .405             | ---        | -              |





|         |   | Data set RN |   |   |
|---------|---|-------------|---|---|
|         |   | Complement  |   |   |
| Control | - | 3           | 0 | 3 |
|         | + | 0           | 2 | 3 |
|         |   | 3           | 2 | 5 |
|         |   | Data set FL |   |   |
|         |   | Complement  |   |   |
| Control | - | 3           | 0 | 3 |
|         | + | 1           | 1 | 2 |
|         |   | 4           | 1 | 5 |
|         |   | Data set SC |   |   |
|         |   | Complement  |   |   |
| Control | - | 1           | 1 | 2 |
|         | + | 2           | 1 | 3 |
|         |   | 3           | 2 | 5 |

Figure 109. 2x2 contingency Tables of reserved data controlled by the anomaly of January rainfall. The complement is the anomaly of the rest-of-year rainfall.



## B. RESULTS

### 1. Data set RN

The model proposed by the maximum likelihood parameters is

$$\theta_t = \frac{e^{.0618+2.9183X_t}}{1 + e^{.0618+2.9183X_t}} \quad \text{VIII.1}$$

This model, when applied to the reserved data yields Table 53.

TABLE 53

RESULTS OF LOGSTIC VALIDATION  
ON DATA SET RN

| YEAR | X      | Y | 0   |
|------|--------|---|-----|
| 1975 | -.739  | 0 | .11 |
| 1976 | -2.746 | 0 | .00 |
| 1977 | -.477  | 0 | .21 |
| 1978 | .883   | 1 | .93 |
| 1979 | .542   | 1 | .84 |
| 1980 | .752   | - | .91 |

The  $\theta_t$  is interpreted, again, as: The conditional probability that the complement, the total rainfall for February through December, will be above its mean value, given that the logged January anomaly was  $X_t$ . Thus it appears that this model tends to predict the direction of the complements deviation well. Figure 110 is a plot of the estimated probabilities against the actual complement anomaly. For an acceptable fit, this plot should show an upward to the right slope, which it does.



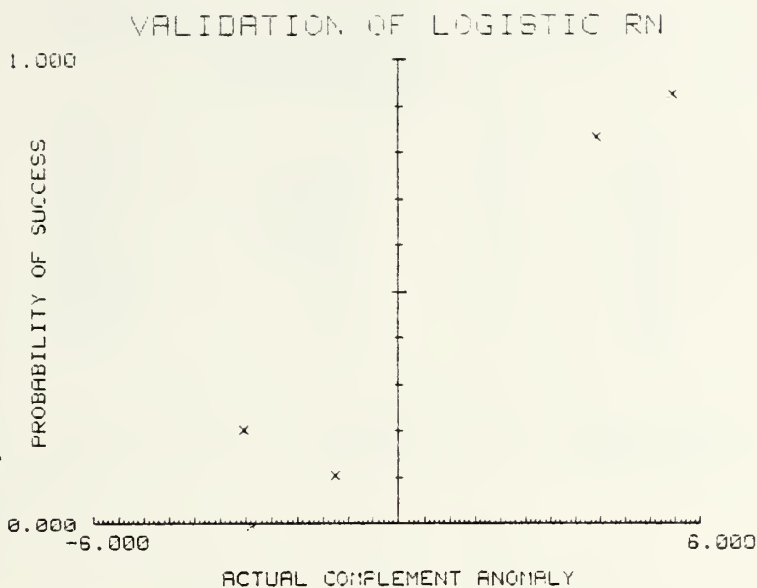


Figure 110. Plot of  $\theta_t$  versus complement anomalies for data set RN

## 2. Data Set FL

This data set is quite similar to the RN data, except that the slope parameter is only a third of that of RN. The model is

$$\theta_t = \frac{e^{-.171+.9325X_t}}{1 + e^{-.171+.9325X_t}} \quad \text{VII.2}$$



TABLE 54

RESULTS OF LOGISTIC VALIDATION  
ON DATA SET FL

| Year | X      | Y | $\theta$ |
|------|--------|---|----------|
| 1975 | -.470  | 0 | .35      |
| 1976 | -2.610 | 0 | .07      |
| 1977 | -.470  | 0 | .35      |
| 1978 | .734   | 1 | .63      |
| 1979 | .547   | 0 | .58      |
| 1980 | .633   | - | .60      |

A plot of the probabilities against the complement anomalies is in Figure 111.

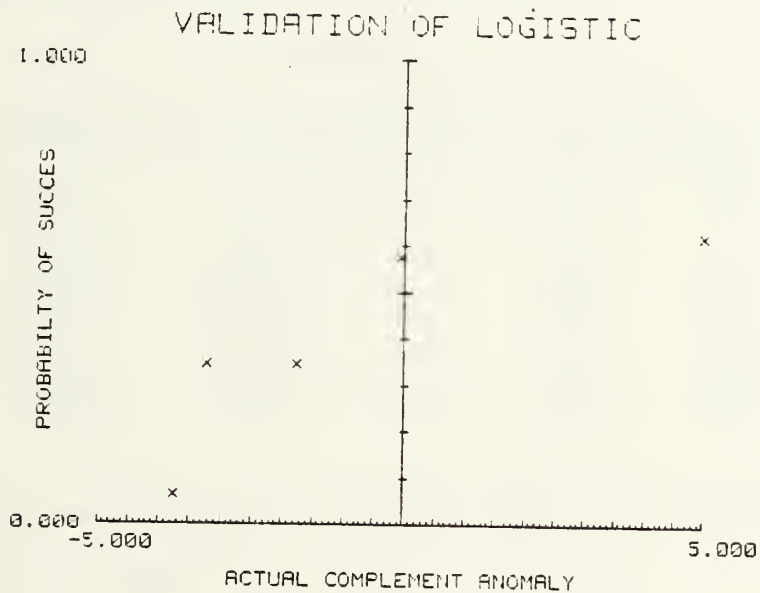


Figure 111. Plot of  $\theta_t$  versus complement anomalies for data set FL

This fit is not as good as that for data set RN. The outlier, or false prediction of 1979 may not, however, be far out of line. The sparsity of points for which the complement anomaly was positive detracts from the validation effort.





### 3. Data Set SC

The maximum likelihood model is

$$\theta_t^{(1)} = \frac{e^{-.3025+.171X_t}}{1 + e^{-.3025+.171X_t}} \quad \text{VII.3}$$

and the IRWLS model is

$$\theta_t^{(2)} = \frac{e^{.1537+.4799X_t}}{1 + e^{.1537+.4799X_t}} \quad \text{VII.4}$$

and the tabular results are in Table 55.

TABLE 55

#### RESULTS OF VALIDATION ON DATA SET SC

| Year | X      | Y | $\theta_t^{(1)}$ | $\theta_t^{(2)}$ |
|------|--------|---|------------------|------------------|
| 1975 | -.535  | 1 | .40              | .47              |
| 1976 | -3.773 | 0 | .28              | .16              |
| 1977 | .352   | 0 | .44              | .58              |
| 1978 | 1.054  | 1 | .47              | .66              |
| 1979 | .512   | 0 | .45              | .60              |
| 1980 | .405   | - | .44              | .59              |

and the plot of the probabilities versus the complement anomalies is in Figure 112.



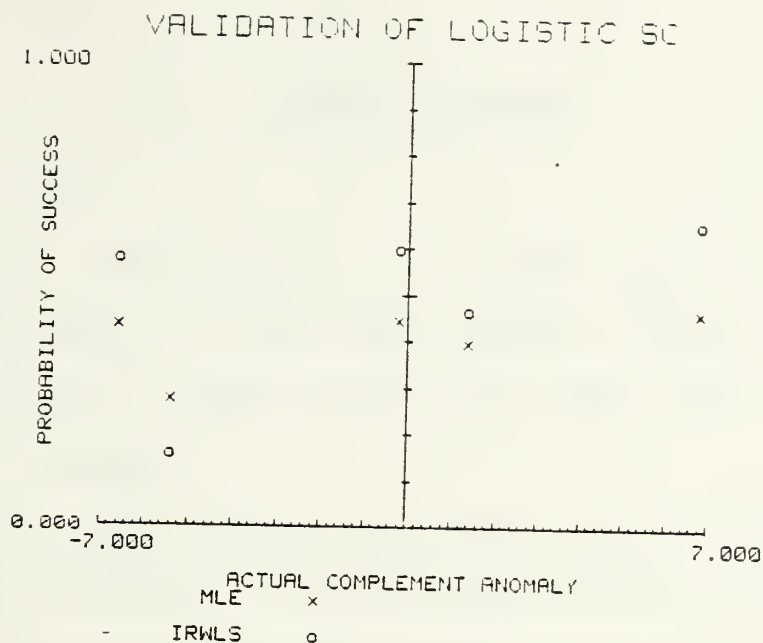


Figure 112. Plot of  $\theta_t$  versus complement anomalies for maximum likelihood and IRWLS parameters for data set SC

### C. DISCUSSION

The validation of the maximum likelihood models for data sets RN and FL appears to be acceptable. However, data set SC does not appear to be acceptably modeled. In fact, as Figure 112 shows, the complement appears to be almost independent of the control. This is also shown by Figure 105 where it can be seen that the contours are very flat and circular about the origin of the  $(\alpha, \beta)$  coordinate system.

The one apparent outlier of data set FL may be viewed as very close, therefore that model can also be assumed to be validated.



## VIII. FURTHER FINDINGS

### A. SUMMER MONTHS

The further investigation of the summer months, to parallel the modeling of the winter months, yielded some interesting results. These results are shown here with no attempt at analysis.

The summer months appear to be increasing in total rainfall and in variance. This is more true for the Peninsula data sets than for the Valley data set. Figures 113 (RN), 114 (FL), and 115 (SC) show the by-month series of summer months. The total summer rainfall series by year are shown in Figures 116 (RN), 117 (FL), and 118 (SC). The reserved data are not included, yet it can be shown to continue the indicated trends.



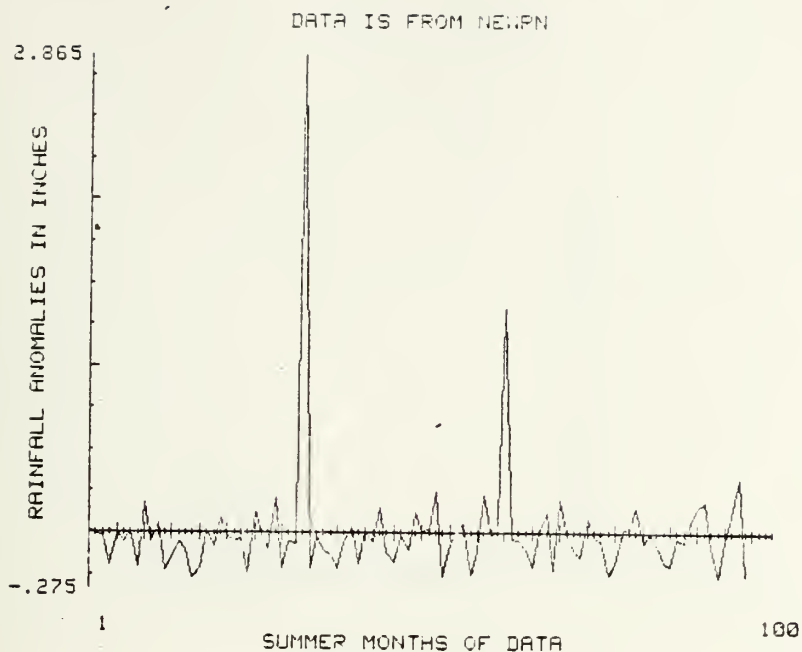


Figure 113. Monthly plot of summer months only, means removed, for data set RN

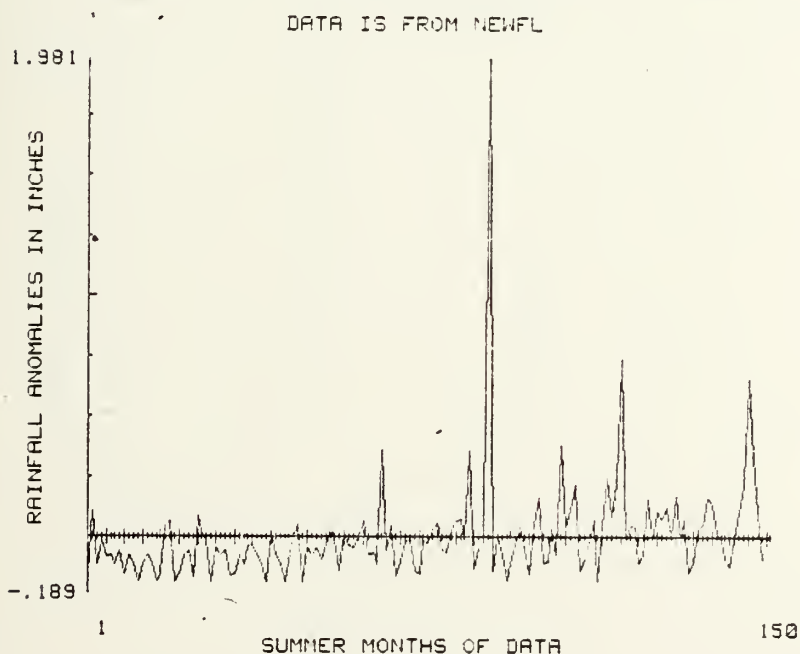


Figure 114. Monthly plot of summer months only, means removed, for data set FL





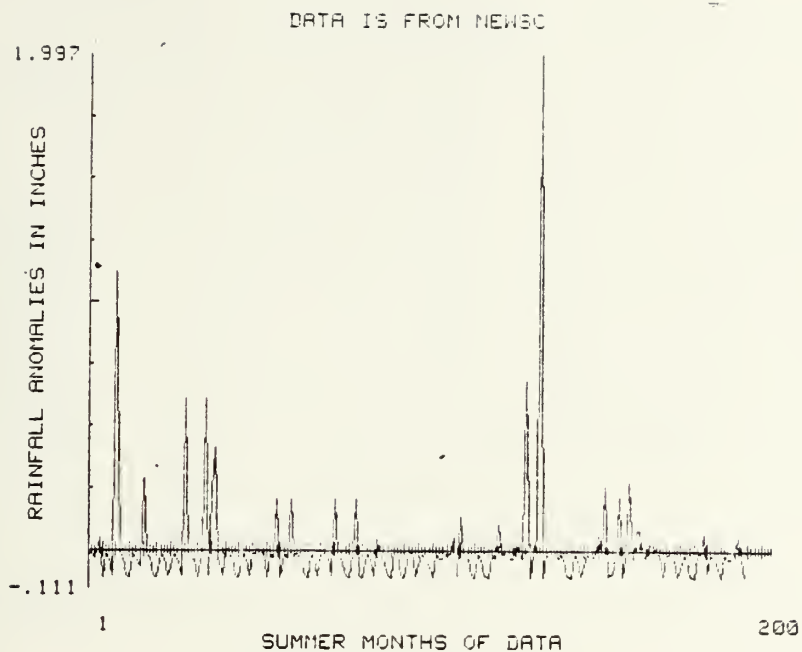


Figure 115. Monthly plot of summer months only, means removed, for data set SC

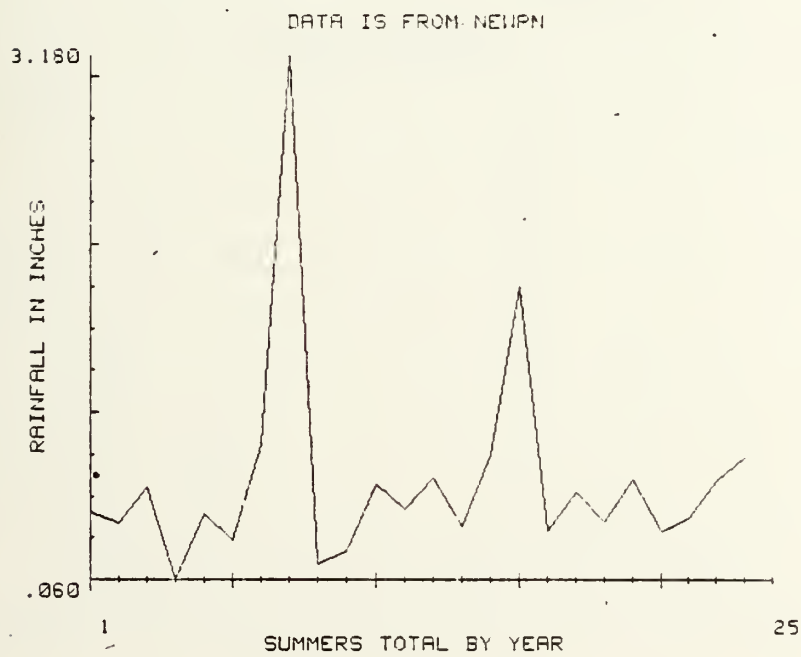


Figure 116. Yearly plot of total summer rainfall for data set RN



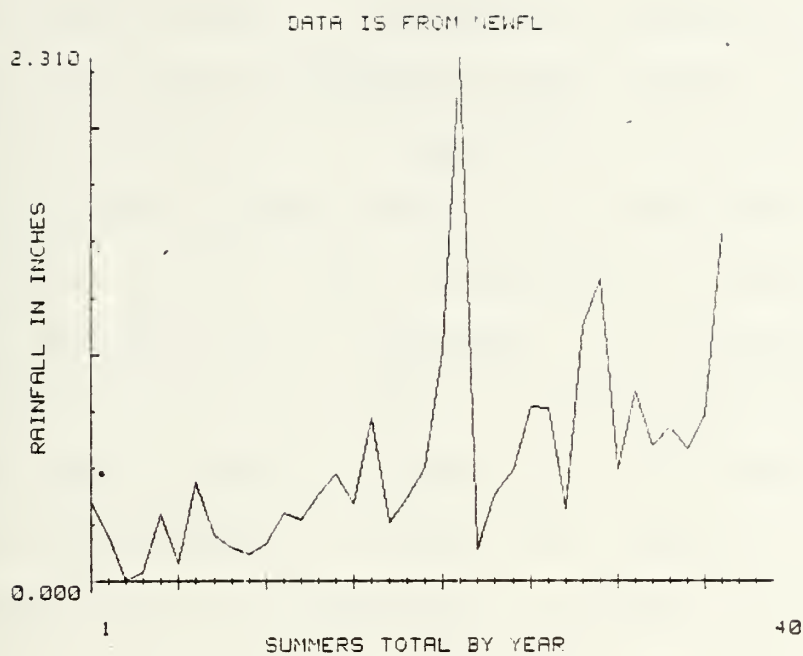


Figure 117. Yearly plot of summer month rainfall for data set FL

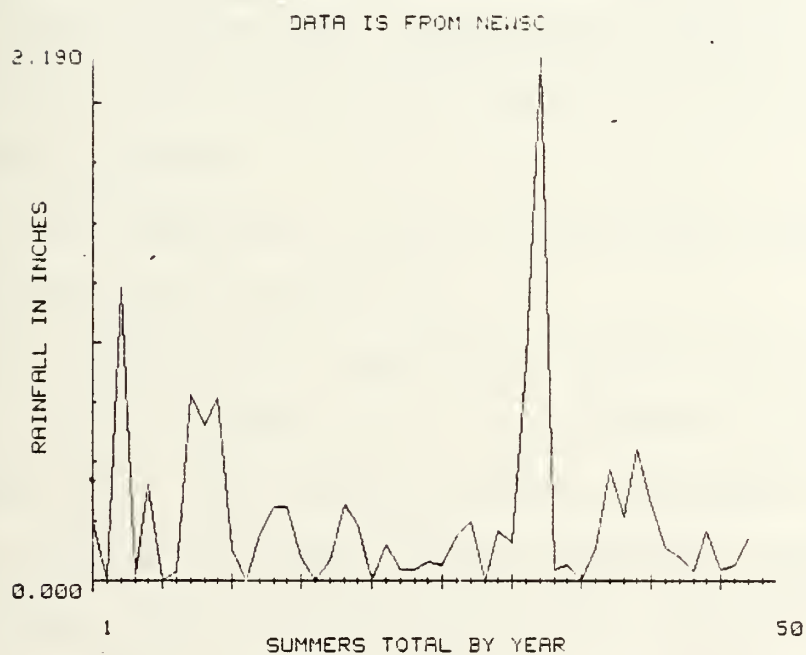


Figure 118. Yearly plot of total summer rainfall for data set SC



## A. SIGNIFICANCE OF JANUARY

The identification of January as a possible predictor for its eleven month complement raises further questions. One of the questions is in determining which part of the complement lends the most towards its predictability. Figure 119 is a plot of the log-odds and chi-square statistics for the cumulative complements. Each progressive column, to the right, of Figure 119 indicates these statistics for another cumulated month, i.e., the first column compares the anomalies of January and February by itself, the second column is a comparison of January to February plus March, and so on until the last column is a comparison of January to the entire eleven month complement.

Several occurrences to be noted from the figure are:

- (1) The log-odds are consistently greater than zero.
- (2) The lack of increased odds and significance during the summer months.
- (3) The similarity of RN to FL and their combined difference to SC in the fall.

These indications suggested a further look at January versus the fall months only. This analysis is displayed in Figure 120. The vertical scales of Figure 119 and Figure 120 are the same, yet the horizontal scales differ. This figure has five major divisions. The left-most division looks at January versus singular months in the fall. The second division looks at January versus pairs



of months in the fall, and so on until the right-most column, which is January versus the total fall rainfall. This figure yields no apparent significance, and unstable odds.

The combined information of Figures 119 and 120 are mildly confusing. One possible explanation may be that the summer months somehow cumulate significance and deviation direction, in order to allow the fall contribution. This possible synergistic affect should be explored further.





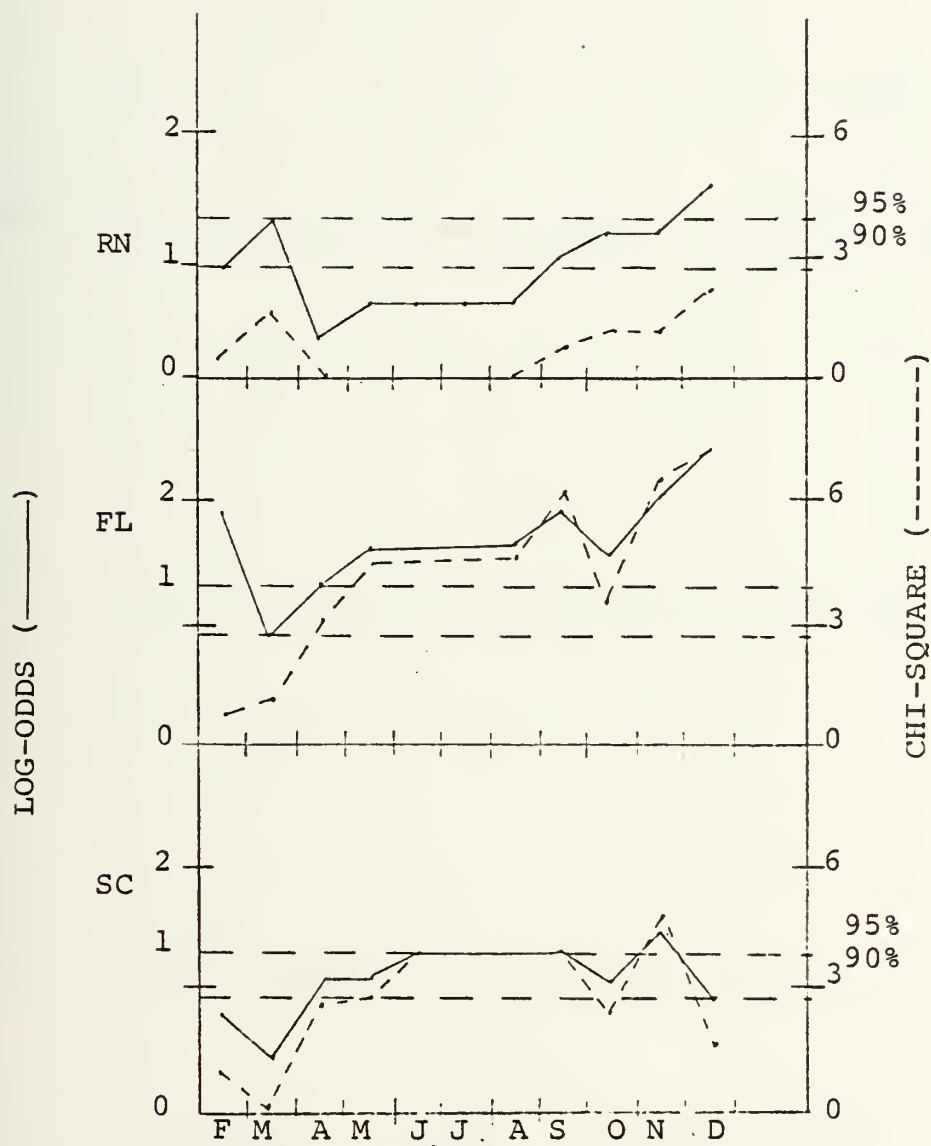


Figure 119. Log-odds and significance versus additional months cumulated through the year



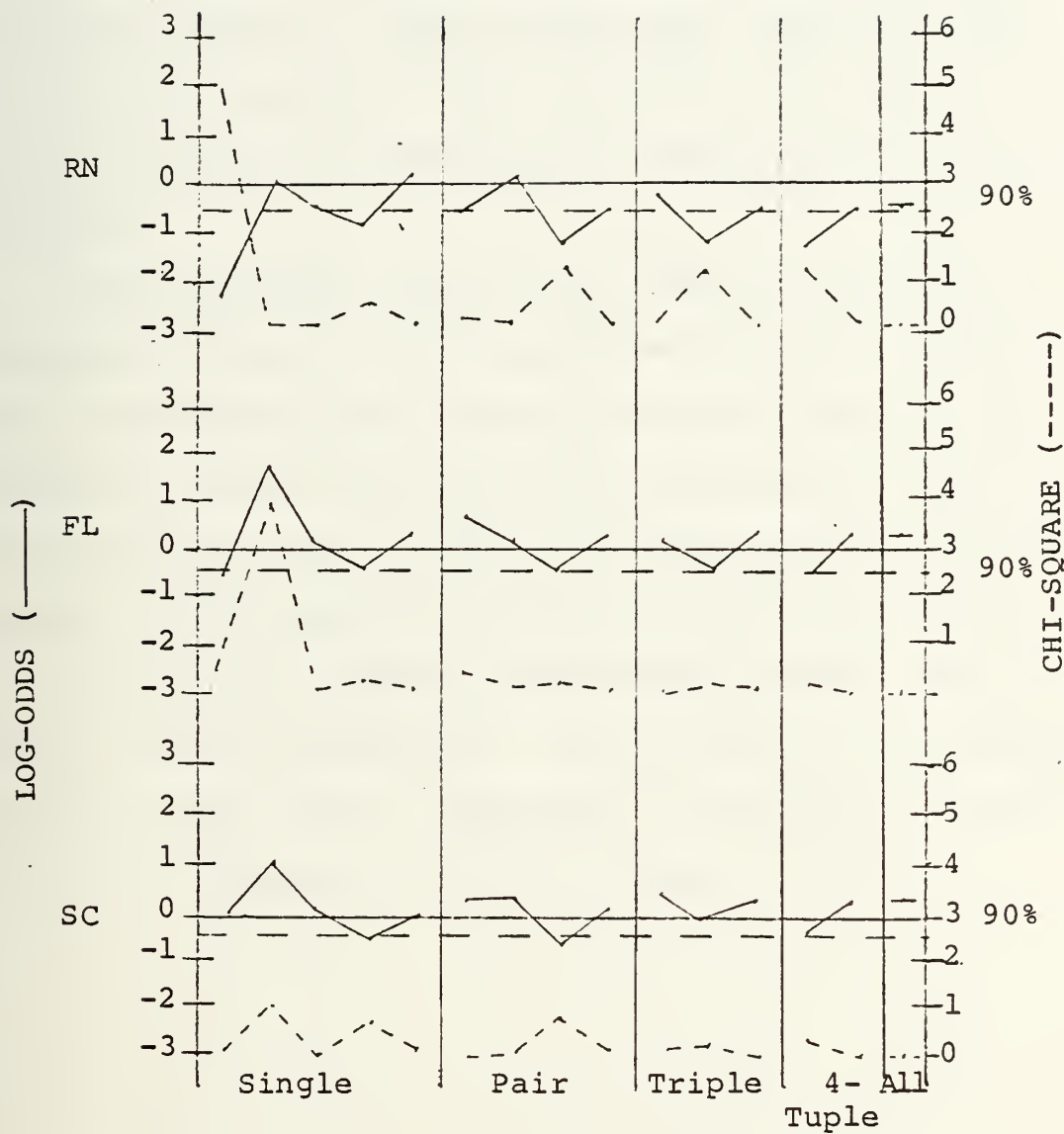


Figure 120. Log-odds and significance versus additional months of the fall



## IX. SUMMARY

The analysis of rainfall data is carried out in a comprehensive way. The autoregressive Markovian model of the early sections could not stand up to validation, but it did point to some sort of dichotomy between the seasons.

2x2 contingency analysis was effective in that it brought attention to the predictive ability of January. This identification of January, when followed by the logistic analysis was seen to be successful in two of the three data sets. Thus, the primary conclusion of this thesis is the predictive ability of January rainfall.

The physical reasoning behind this finding must be left to the meteorologist. Further study of the approach used here may lead to improvement in seasonal or annual rainfall forecasts for certain climatic regions.



## APPENDIX A

## DATA SET RN

RAW

| YEAR | OCT<br>APR     | NOV<br>MAY     | DEC<br>JUN     | JAN<br>JUL     | FEB<br>AUG     | MAR<br>SEP    |
|------|----------------|----------------|----------------|----------------|----------------|---------------|
| 1951 | .900<br>1.100  | 2.780<br>.160  | 6.880<br>.240  | 10.040<br>.060 | 2.960<br>.080  | 4.410<br>.080 |
| 1952 | .170<br>1.670  | 2.190<br>.490  | 5.900<br>.220  | 2.060<br>0.000 | .040<br>.110   | 1.200<br>.070 |
| 1953 | .380<br>.850   | 2.150<br>.420  | .560<br>.400   | 4.260<br>0.000 | 2.260<br>.160  | 4.910<br>.050 |
| 1954 | .030<br>1.740  | 2.500<br>.870  | 3.130<br>.060  | 5.820<br>0.000 | 1.740<br>0.000 | .160<br>0.000 |
| 1955 | .080<br>1.670  | 1.960<br>.590  | 9.790<br>0.000 | 6.090<br>.070  | .060<br>.020   | .150<br>.360  |
| 1956 | 1.000<br>1.490 | 0.000<br>2.390 | .840<br>.200   | 4.650<br>0.000 | 3.520<br>.070  | 1.920<br>.030 |
| 1957 | 1.580<br>4.710 | .930<br>.560   | 3.700<br>.350  | 3.710<br>.040  | 5.660<br>0.000 | 7.170<br>.480 |
| 1958 | .040<br>.290   | .510<br>.120   | .490<br>0.000  | 4.850<br>0.000 | 5.760<br>.040  | .320<br>3.140 |
| 1959 | 0.000<br>.880  | 0.000<br>.340  | .590<br>0.000  | 4.300<br>.030  | 4.530<br>0.000 | .840<br>.130  |
| 1960 | .070<br>1.290  | 2.060<br>.720  | .850<br>0.000  | 1.890<br>0.000 | 1.170<br>.140  | 2.580<br>.090 |
| 1961 | .040<br>.300   | 1.740<br>.150  | 1.190<br>.230  | 2.690<br>0.000 | 5.170<br>.250  | 2.570<br>.150 |
| 1962 | 1.330<br>3.930 | .370<br>.660   | 2.210<br>.040  | 3.050<br>.040  | 2.700<br>.001  | 4.140<br>.400 |
| 1963 | 1.460<br>.220  | 3.770<br>.860  | .530<br>.220   | 3.500<br>.090  | .420<br>.350   | 2.230<br>.010 |
| 1964 | .780<br>2.260  | 3.290<br>.170  | 6.450<br>.150  | 2.560<br>.050  | 1.050<br>.160  | 2.440<br>.020 |
| 1965 | .230<br>.270   | 6.490<br>.130  | 5.560<br>.120  | 2.320<br>.280  | 1.880<br>.090  | .430<br>.320  |
| 1966 | .090<br>7.110  | 4.740<br>.400  | 4.180<br>1.560 | 5.290<br>.020  | .450<br>.060   | 5.480<br>.170 |





| YEAR | OCT<br>APR     | NOV<br>MAY    | DEC<br>JUN    | JAN<br>JUL    | FEB<br>AUG    | MAR<br>SEP    |
|------|----------------|---------------|---------------|---------------|---------------|---------------|
| 1967 | .380<br>.790   | 1.610<br>.320 | 2.270<br>.010 | 3.100<br>.060 | 1.400<br>.230 | 3.050<br>.050 |
| 1968 | .310<br>2.700  | 3.130<br>.120 | 3.270<br>.420 | 9.450<br>.040 | 7.310<br>.001 | 1.310<br>.120 |
| 1969 | .500<br>.350   | .720<br>.050  | 3.080<br>.300 | 5.910<br>.030 | 2.040<br>.060 | 2.970<br>.020 |
| 1970 | .590<br>1.190  | 6.170<br>.710 | 4.990<br>.030 | 1.080<br>.070 | .620<br>.130  | 1.960<br>.430 |
| 1971 | .090<br>.880   | 1.990<br>.090 | 4.760<br>.150 | 1.230<br>.060 | 1.050<br>.040 | .030<br>.100  |
| 1972 | 2.460<br>.130  | 5.950<br>.060 | 2.080<br>.020 | 6.050<br>.020 | 5.880<br>.050 | 4.520<br>.340 |
| 1973 | 2.200<br>3.400 | 3.870<br>.030 | 4.730<br>.370 | 3.730<br>.250 | .910<br>.020  | 4.480<br>.010 |
| 1974 | 1.540<br>1.760 | .560<br>.010  | 2.480<br>.170 | 1.340<br>.170 | 3.620<br>.430 | 4.060<br>.020 |



# Logged Anomalies

| YEAR | OCT<br>APR     | NOV<br>MAY      | DEC<br>JUN     | JAN<br>JUL     | FEB<br>AUG      | MAR<br>SEP     |
|------|----------------|-----------------|----------------|----------------|-----------------|----------------|
| 1951 | .204<br>-.112  | .237<br>-.179   | .753<br>.040   | .863<br>.005   | .272<br>-.017   | .556<br>-.108  |
| 1952 | -.291<br>.128  | .068<br>.080    | .620<br>.023   | -.421<br>-.054 | -1.065<br>.010  | -.344<br>-.118 |
| 1953 | -.116<br>-.239 | .055<br>.032    | -.867<br>.161  | .121<br>-.054  | .078<br>.054    | .644<br>-.137  |
| 1954 | -.409<br>.154  | .160<br>.307    | .106<br>-.117  | .381<br>-.054  | -.096<br>-.094  | -.984<br>-.185 |
| 1955 | -.361<br>.128  | -.007<br>.145   | 1.067<br>-.176 | .420<br>.014   | -1.046<br>-.074 | -.993<br>.122  |
| 1956 | .255<br>.058   | -1.092<br>.902  | -.702<br>.007  | .193<br>-.054  | .404<br>-.026   | -.061<br>-.156 |
| 1957 | .510<br>.888   | -.435<br>.126   | .236<br>.125   | .011<br>-.014  | .792<br>-.094   | .968<br>.207   |
| 1958 | -.399<br>-.599 | -.680<br>-.205  | -.913<br>-.176 | .228<br>-.054  | .807<br>-.055   | -.855<br>1.235 |
| 1959 | -.438<br>-.223 | -1.092<br>-.026 | -.848<br>-.176 | .129<br>-.024  | .606<br>-.094   | -.523<br>-.063 |
| 1960 | -.370<br>-.025 | .026<br>.224    | -.697<br>-.176 | -.478<br>-.054 | -.329<br>.037   | .143<br>-.099  |
| 1961 | -.399<br>-.592 | -.084<br>-.179  | -.528<br>.031  | -.233<br>-.054 | .715<br>.129    | .140<br>-.046  |
| 1962 | .408<br>.741   | -.778<br>.188   | -.146<br>-.136 | -.140<br>-.014 | .204<br>-.093   | .504<br>.151   |
| 1963 | .462<br>-.655  | .470<br>.302    | -.887<br>.023  | -.035<br>.032  | -.754<br>.206   | .040<br>-.175  |
| 1964 | .139<br>.328   | .364<br>-.162   | .696<br>-.036  | -.269<br>-.005 | -.386<br>.054   | .103<br>-.166  |
| 1965 | -.231<br>-.615 | .921<br>-.196   | .569<br>-.062  | -.339<br>.193  | -.046<br>-.008  | -.775<br>.092  |
| 1966 | -.352<br>1.239 | .655<br>.018    | .333<br>.764   | .300<br>-.034  | -.733<br>-.036  | .736<br>-.028  |



| YEAR | OCT<br>APR     | NOV<br>MAY     | DEC<br>JUN     | JAN<br>JUL    | FEB<br>AUG     | MAR<br>SEP      |
|------|----------------|----------------|----------------|---------------|----------------|-----------------|
| 1967 | -.116<br>-.272 | -.133<br>-.041 | -.127<br>-.166 | -.123<br>.005 | -.229<br>.113  | .269<br>-.137   |
| 1968 | -.168<br>.454  | .326<br>-.205  | .140<br>.175   | .808<br>-.014 | 1.013<br>-.093 | -.295<br>-.072  |
| 1969 | -.033<br>-.554 | -.550<br>-.270 | .094<br>.097   | .394<br>-.024 | .008<br>-.036  | .246<br>-.166   |
| 1970 | .026<br>-.070  | .878<br>.218   | .478<br>-.146  | -.807<br>.014 | -.622<br>.028  | -.047<br>.172   |
| 1971 | -.352<br>-.223 | .003<br>-.232  | .439<br>-.036  | -.737<br>.005 | -.386<br>-.055 | -1.103<br>-.090 |
| 1972 | .803<br>-.732  | .846<br>-.260  | -.187<br>-.156 | .414<br>-.034 | .824<br>-.045  | .576<br>.107    |
| 1973 | .725<br>.628   | .491<br>-.289  | .434<br>.139   | .015<br>.169  | -.457<br>-.074 | .568<br>-.175   |
| 1974 | .494<br>.161   | -.648<br>-.309 | -.065<br>-.019 | -.689<br>.103 | .426<br>.264   | .489<br>-.166   |



# Reserved Data

| YEAR | OCT<br>APR     | NOV<br>MAY    | DEC<br>JUN    | JAN<br>JUL    | FEB<br>AUG    | MAR<br>SEP    |
|------|----------------|---------------|---------------|---------------|---------------|---------------|
| 1974 | 1.540<br>1.760 | .560<br>.010  | 2.480<br>.170 | 1.340<br>.170 | 3.620<br>.430 | 4.060<br>.020 |
| 1975 | 1.700<br>1.740 | .520<br>.070  | .370<br>.170  | .180<br>.020  | 2.970<br>.970 | 1.520<br>.420 |
| 1976 | .600<br>.040   | .720<br>1.210 | 2.080<br>.080 | 1.740<br>.030 | .830<br>.020  | 1.750<br>.650 |
| 1977 | .140<br>5.430  | .540<br>.020  | 5.850<br>.080 | 6.780<br>.040 | 4.780<br>.001 | 5.240<br>.290 |
| 1978 | .020<br>.580   | 2.130<br>.290 | 1.590<br>.020 | 4.820<br>.350 | 4.520<br>.090 | 4.410<br>.020 |
| 1979 | 1.800<br>1.770 | 2.850<br>.570 | 3.180<br>.040 | 5.950<br>.730 | 4.780<br>.090 | 2.400<br>.090 |





APPENDIX B  
DATA SET FL

RAW

| YEAR | OCT<br>APR     | NOV<br>MAY    | DEC<br>JUN     | JAN<br>JUL     | FEB<br>AUG     | MAR<br>SEP     |
|------|----------------|---------------|----------------|----------------|----------------|----------------|
| 1937 | .640<br>2.250  | 1.270<br>.050 | 4.600<br>0.000 | 4.990<br>.180  | 7.590<br>0.000 | 5.630<br>.170  |
| 1938 | 1.280<br>.280  | .770<br>.630  | 2.090<br>.070  | 3.190<br>0.000 | 1.620<br>0.000 | 2.370<br>.130  |
| 1939 | .750<br>.720   | .480<br>.130  | 1.470<br>0.000 | 0.430<br>0.000 | 8.200<br>0.000 | 4.240<br>0.000 |
| 1940 | 0.000<br>5.760 | .790<br>.470  | 4.520<br>.040  | 6.150<br>0.000 | 7.730<br>0.000 | 5.500<br>0.000 |
| 1941 | .800<br>3.660  | .280<br>.790  | 7.870<br>0.000 | 3.480<br>.120  | 2.910<br>.180  | 2.430<br>0.000 |
| 1942 | 1.130<br>1.300 | 1.470<br>.100 | .970<br>0.000  | 3.050<br>0.000 | 2.750<br>.060  | 4.240<br>.020  |
| 1943 | .620<br>1.160  | .420<br>.830  | 3.830<br>.240  | 3.690<br>.100  | 8.080<br>.100  | .710<br>0.000  |
| 1944 | 1.130<br>.430  | 4.650<br>.300 | 1.940<br>.110  | .980<br>0.000  | 3.180<br>.060  | 2.830<br>.030  |
| 1945 | 1.370<br>.040  | 1.400<br>.640 | 4.230<br>0.000 | 1.080<br>0.000 | 2.560<br>0.000 | 3.000<br>.150  |
| 1946 | .260<br>.330   | 4.310<br>.210 | 1.970<br>.120  | .440<br>0.000  | 1.130<br>0.000 | 1.950<br>0.000 |
| 1947 | 1.160<br>3.250 | .470<br>.400  | 1.940<br>.170  | .090<br>0.000  | 2.140<br>0.000 | 4.610<br>0.000 |
| 1948 | 2.130<br>.080  | .350<br>.420  | 3.200<br>.050  | 1.700<br>.080  | 2.840<br>.170  | 4.730<br>0.000 |
| 1949 | .030<br>1.470  | 1.690<br>.190 | 1.370<br>.110  | 3.120<br>0.000 | 1.740<br>.070  | 1.640<br>.090  |
| 1950 | 1.760<br>1.060 | 3.000<br>.270 | 2.370<br>.120  | 2.060<br>.110  | 2.190<br>.100  | 1.480<br>.050  |
| 1951 | .670<br>1.100  | 3.350<br>.120 | 6.040<br>.170  | 8.010<br>.040  | 2.110<br>.070  | 4.550<br>.190  |
| 1952 | .140<br>1.250  | 2.210<br>.570 | 5.040<br>.220  | 1.710<br>0.000 | 0.000<br>.050  | .770<br>.070   |



| YEAR | OCT<br>APR     | NOV<br>MAY     | DEC<br>JUN     | JAN<br>JUL     | FEB<br>AUG    | MAR<br>SEP    |
|------|----------------|----------------|----------------|----------------|---------------|---------------|
| 1953 | .280<br>.780   | 2.050<br>.390  | .310<br>.510   | 3.330<br>.020  | 1.580<br>.160 | 4.410<br>.030 |
| 1954 | .020<br>1.660  | 2.170<br>.960  | 2.820<br>.050  | 4.560<br>.060  | 1.560<br>.100 | .130<br>.050  |
| 1955 | .040<br>1.560  | 1.650<br>.590  | 8.310<br>0.000 | 4.820<br>.060  | 1.740<br>.090 | .180<br>.220  |
| 1956 | .650<br>1.340  | 0.000<br>2.640 | .770<br>.210   | 4.320<br>.020  | 2.610<br>.050 | 1.890<br>.220 |
| 1957 | 2.020<br>3.980 | .850<br>.490   | 3.380<br>.220  | 2.620<br>.150  | 4.620<br>.100 | 6.760<br>.550 |
| 1958 | .040<br>.460   | .390<br>.120   | .400<br>.020   | 4.470<br>.030  | 5.630<br>.090 | .340<br>2.170 |
| 1959 | .030<br>.870   | 0.000<br>.410  | .620<br>.010   | 3.990<br>.080  | 3.270<br>.040 | .650<br>.010  |
| 1960 | .070<br>1.200  | 1.450<br>.750  | .640<br>.040   | 1.840<br>.040  | 1.110<br>.160 | 2.420<br>.140 |
| 1961 | .100<br>.270   | 2.030<br>.150  | 1.380<br>0.000 | 2.230<br>.130  | 5.830<br>.280 | 2.400<br>.080 |
| 1962 | 1.360<br>3.320 | .440<br>.650   | 2.260<br>.050  | 3.830<br>.110  | 2.270<br>.040 | 4.320<br>.570 |
| 1963 | 1.360<br>.100  | 5.260<br>.690  | .540<br>.190   | 3.190<br>.190  | .390<br>.330  | 2.300<br>.050 |
| 1964 | 1.060<br>2.180 | 3.160<br>.140  | 5.970<br>.060  | 2.370<br>.060  | .930<br>.190  | 2.430<br>.010 |
| 1965 | .160<br>.170   | 6.250<br>.140  | 6.230<br>.190  | 3.050<br>.320  | 1.490<br>.200 | .510<br>.410  |
| 1966 | .140<br>6.810  | 4.690<br>.790  | 3.880<br>.890  | 5.750<br>.060  | .490<br>.160  | 5.070<br>.230 |
| 1967 | .270<br>.720   | 1.810<br>.290  | 2.060<br>.040  | 3.110<br>0.000 | 1.350<br>.270 | 3.180<br>.180 |
| 1968 | .430<br>2.490  | 3.300<br>.190  | 2.750<br>.260  | 8.480<br>.150  | 7.960<br>.240 | 1.200<br>.190 |



| YEAR | OCT<br>APR     | NOV<br>MAY    | DEC<br>JUN    | JAN<br>JUL    | FEB<br>AUG    | MAR<br>SEP    |
|------|----------------|---------------|---------------|---------------|---------------|---------------|
| 1969 | .410<br>.430   | 1.010<br>.080 | 2.800<br>.320 | 5.550<br>.050 | 3.020<br>.190 | 1.850<br>.040 |
| 1970 | .540<br>1.190  | 7.440<br>.580 | 4.450<br>.050 | .870<br>.110  | .640<br>.170  | 1.940<br>.350 |
| 1971 | .170<br>.720   | 1.550<br>.180 | 4.680<br>.290 | 1.000<br>.090 | .440<br>.100  | .080<br>.100  |
| 1972 | 2.540<br>.310  | 5.520<br>.160 | 1.850<br>.030 | 5.550<br>.070 | 6.070<br>.220 | 3.530<br>.410 |
| 1973 | 1.960<br>2.990 | 4.750<br>.110 | 3.270<br>.810 | 3.760<br>.410 | 1.190<br>.210 | 4.500<br>.100 |



# Logged Anomalies

| YEAR | OCT<br>APR     | NOV<br>MAY     | DEC<br>JUN     | JAN<br>JUL      | FEB<br>AUG      | MAR<br>SEP     |
|------|----------------|----------------|----------------|-----------------|-----------------|----------------|
| 1937 | .011<br>.383   | -.187<br>-.289 | .454<br>-.130  | .392<br>.095    | .932<br>-.106   | .708<br>.011   |
| 1938 | .340<br>-.549  | -.436<br>.151  | -.141<br>-.062 | .035<br>-.071   | -.255<br>-.106  | .031<br>-.024  |
| 1939 | .076<br>-.254  | -.615<br>-.216 | -.365<br>-.130 | .846<br>-.071   | 1.001<br>-.106  | .473<br>-.146  |
| 1940 | -.484<br>1.115 | -.425<br>.047  | .439<br>-.091  | .569<br>-.071   | .948<br>-.106   | .688<br>-.146  |
| 1941 | .104<br>.743   | -.760<br>.244  | .914<br>-.130  | .102<br>.042    | .145<br>.060    | .049<br>-.146  |
| 1942 | .272<br>.037   | -.103<br>-.243 | -.591<br>-.130 | .001<br>-.071   | .103<br>-.047   | .473<br>-.126  |
| 1943 | -.001<br>-.026 | -.656<br>.266  | .306<br>.085   | .148<br>.024    | .988<br>-.010   | -.647<br>-.146 |
| 1944 | .272<br>-.438  | .725<br>-.076  | -.191<br>-.025 | -.715<br>-.071  | .212<br>-.047   | .159<br>-.117  |
| 1945 | .379<br>-.757  | -.132<br>.157  | .385<br>-.130  | -.666<br>-.071  | .051<br>-.106   | .203<br>-.006  |
| 1946 | -.253<br>-.511 | .663<br>-.147  | -.181<br>-.017 | -1.033<br>-.071 | -.462<br>-.106  | -.102<br>-.146 |
| 1947 | .286<br>.651   | -.622<br>-.001 | -.191<br>.027  | -1.312<br>-.071 | -.074<br>-.106  | .541<br>-.146  |
| 1948 | .657<br>-.719  | -.707<br>.013  | .166<br>-.081  | -.405<br>.006   | .127<br>.051    | .562<br>-.146  |
| 1949 | -.454<br>.108  | -.018<br>-.164 | -.406<br>-.025 | .018<br>-.071   | -.210<br>-.038  | -.213<br>-.060 |
| 1950 | .532<br>-.073  | .379<br>-.099  | -.054<br>-.017 | -.279<br>.034   | -.058<br>-.010  | -.275<br>-.097 |
| 1951 | .029<br>-.054  | .463<br>-.225  | .682<br>.027   | .800<br>-.032   | -.084<br>-.038  | .530<br>.028   |
| 1952 | -.353<br>.015  | .159<br>.113   | .529<br>.069   | -.401<br>-.071  | -1.218<br>-.057 | -.613<br>-.078 |





| YEAR | OCT<br>APR     | NOV<br>MAY     | DEC<br>JUN     | JAN<br>JUL     | FEB<br>AUG     | MAR<br>SEP      |
|------|----------------|----------------|----------------|----------------|----------------|-----------------|
| 1953 | -.237<br>-.219 | .108<br>-.009  | -.999<br>.232  | .063<br>-.051  | -.271<br>.043  | .505<br>-.117   |
| 1954 | -.464<br>.182  | .147<br>.335   | .071<br>-.081  | .318<br>-.013  | -.278<br>-.010 | -1.061<br>-.097 |
| 1955 | -.444<br>.144  | -.032<br>.126  | .962<br>-.130  | .363<br>-.013  | -.210<br>-.019 | -1.018<br>.053  |
| 1956 | .017<br>.054   | -1.007<br>.954 | -.698<br>.061  | .274<br>-.051  | .065<br>-.057  | -.122<br>.053   |
| 1957 | .622<br>.810   | -.392<br>.061  | .208<br>.069   | -.111<br>.069  | .508<br>-.010  | .865<br>.292    |
| 1958 | -.444<br>-.417 | -.673<br>-.225 | -.933<br>-.110 | .301<br>-.041  | .673<br>-.019  | -.891<br>1.008  |
| 1959 | -.454<br>-.170 | -1.007<br>.006 | -.787<br>-.120 | .210<br>.006   | .233<br>-.066  | -.683<br>-.136  |
| 1960 | -.416<br>-.007 | -.111<br>.222  | -.774<br>-.091 | -.354<br>-.032 | -.472<br>.043  | .046<br>-.015   |
| 1961 | -.388<br>-.557 | .102<br>-.198  | -.402<br>-.130 | -.225<br>.051  | .703<br>.141   | .040<br>-.069   |
| 1962 | .375<br>.667   | -.642<br>.163  | -.087<br>-.081 | .177<br>.034   | -.034<br>-.066 | .488<br>.305    |
| 1963 | .375<br>-.701  | .827<br>.187   | -.837<br>.044  | .035<br>.103   | -.889<br>.180  | .010<br>-.097   |
| 1964 | .239<br>.361   | .418<br>-.207  | .673<br>-.072  | -.183<br>-.013 | -.561<br>.068  | .049<br>-.136   |
| 1965 | -.335<br>-.639 | .974<br>-.207  | .709<br>.044   | .001<br>.207   | -.306<br>.077  | -.771<br>.198   |
| 1966 | -.353<br>1.260 | .732<br>.244   | .316<br>.507   | .512<br>-.013  | -.820<br>.043  | .620<br>.061    |
| 1967 | -.245<br>-.254 | .026<br>-.083  | -.151<br>-.091 | .016<br>-.071  | -.364<br>.133  | .247<br>.019    |
| 1968 | -.126<br>.454  | .452<br>-.164  | .053<br>.101   | .851<br>.069   | .974<br>.110   | -.395<br>.028   |



| YEAR | OCT<br>APR     | NOV<br>MAY     | DEC<br>JUN     | JAN<br>JUL    | FEB<br>AUG     | MAR<br>SEP      |
|------|----------------|----------------|----------------|---------------|----------------|-----------------|
| 1969 | -.140<br>-.438 | -.309<br>-.261 | .066<br>.148   | .482<br>-.022 | .173<br>.068   | -.133<br>-.107  |
| 1970 | -.052<br>-.012 | 1.126<br>.120  | .427<br>-.081  | -.772<br>.034 | -.724<br>.051  | -.105<br>.154   |
| 1971 | -.327<br>-.254 | -.071<br>-.172 | .468<br>.125   | -.705<br>.015 | -.854<br>-.010 | -1.107<br>-.051 |
| 1972 | .780<br>-.526  | .868<br>-.189  | -.222<br>-.100 | .482<br>-.003 | .738<br>.093   | .327<br>.198    |
| 1973 | .602<br>.588   | .742<br>-.234  | .183<br>.463   | .162<br>.273  | -.434<br>.085  | .521<br>-.051   |



# Reserved Data

| YEAR | OCT<br>APR     | NOV<br>MAY    | DEC<br>JUN    | JAN<br>JUL    | FEB<br>AUG     | MAR<br>SEP     |
|------|----------------|---------------|---------------|---------------|----------------|----------------|
| 1974 | 1.960<br>1.580 | .720<br>.110  | 2.430<br>.180 | 1.700<br>.380 | 2.760<br>.450  | 4.680<br>.080  |
| 1975 | 1.480<br>1.660 | .470<br>.070  | .410<br>.130  | .200<br>.140  | 2.820<br>1.330 | 1.060<br>.820  |
| 1976 | .110<br>.030   | .780<br>.960  | 1.660<br>.110 | 1.700<br>.020 | .630<br>.040   | 1.900<br>1.050 |
| 1977 | .170<br>5.210  | .620<br>.030  | 5.570<br>.100 | 5.780<br>.090 | 4.710<br>.050  | 5.160<br>.310  |
| 1978 | .100<br>.560   | 2.220<br>.180 | 1.340<br>.010 | 4.700<br>.330 | 3.650<br>.120  | 3.710<br>.040  |
| 1979 | 1.580<br>1.520 | 3.360<br>.580 | 3.830<br>.040 | 5.120<br>.900 | 4.420<br>.140  | 2.030<br>.130  |



## APPENDIX C

## DATA SET SC

RAW

| YEAR | OCT<br>APR     | NOV<br>MAY     | DEC<br>JUN      | JAN<br>JUL     | FEB<br>AUG      | MAR<br>SEP     |
|------|----------------|----------------|-----------------|----------------|-----------------|----------------|
| 1926 | .420<br>1.070  | 3.360<br>.280  | 1.640<br>.100   | 2.440<br>0.000 | 7.530<br>0.000  | 1.390<br>.150  |
| 1927 | .800<br>.960   | 2.100<br>.050  | 2.330<br>0.000  | .830<br>0.000  | 2.430<br>0.000  | 3.580<br>0.000 |
| 1928 | .020<br>1.290  | 3.120<br>0.000 | 3.410<br>1.230  | 1.260<br>0.000 | 1.370<br>0.000  | 2.890<br>0.000 |
| 1929 | 0.000<br>1.330 | 0.000<br>1.610 | .330<br>0.000   | 5.370<br>0.000 | 3.260<br>0.000  | 4.250<br>.030  |
| 1930 | .050<br>.490   | 1.480<br>.660  | .050<br>.400    | 4.260<br>0.000 | 1.560<br>0.000  | .920<br>0.000  |
| 1931 | .020<br>.140   | 1.750<br>.360  | 10.260<br>0.000 | 4.320<br>0.000 | 4.690<br>0.000  | .800<br>0.000  |
| 1932 | 0.000<br>.330  | .180<br>.320   | 3.120<br>.040   | 6.920<br>0.000 | .920<br>0.000   | 1.570<br>0.000 |
| 1933 | 1.120<br>.130  | 0.000<br>.850  | 3.250<br>.720   | 3.150<br>0.000 | 5.290<br>0.000  | 0.000<br>.060  |
| 1934 | .150<br>5.910  | 2.890<br>0.000 | 2.690<br>0.000  | 7.000<br>0.000 | .770<br>.650    | 4.880<br>0.000 |
| 1935 | .370<br>2.460  | .810<br>.480   | 2.010<br>.340   | 2.660<br>.430  | 9.970<br>0.000  | 1.630<br>0.000 |
| 1936 | .430<br>.520   | 0.000<br>0.000 | 4.160<br>.130   | 4.620<br>0.000 | 6.770<br>0.000  | 7.730<br>0.000 |
| 1937 | .060<br>2.630  | 1.010<br>0.000 | 6.590<br>0.000  | 3.440<br>0.000 | 13.020<br>0.000 | 3.090<br>0.000 |
| 1938 | .810<br>.420   | 1.040<br>.220  | 2.430<br>.100   | 3.540<br>0.000 | 2.700<br>0.000  | 3.030<br>.100  |
| 1939 | 1.040<br>.600  | .370<br>.270   | 2.330<br>0.000  | 9.200<br>0.000 | 3.960<br>0.000  | 2.320<br>.310  |
| 1940 | 1.040<br>.600  | .370<br>.270   | 2.330<br>0.000  | 9.200<br>0.000 | 3.960<br>0.000  | 2.320<br>.310  |
| 1941 | .450<br>5.080  | .330<br>.610   | 9.450<br>.100   | 5.130<br>0.000 | 10.420<br>0.000 | 3.920<br>0.000 |





| YEAR | OCT<br>APR     | NOV<br>MAY     | DEC<br>JUN      | JAN<br>JUL     | FEB<br>AUG     | MAR<br>SEP     |
|------|----------------|----------------|-----------------|----------------|----------------|----------------|
| 1942 | 1.130<br>4.910 | 1.040<br>1.010 | 10.490<br>0.000 | 5.710<br>0.000 | 2.460<br>0.000 | 3.280<br>0.000 |
| 1943 | .740<br>1.070  | 1.960<br>0.000 | 1.630<br>.090   | 8.400<br>0.000 | 2.420<br>0.000 | 3.740<br>0.000 |
| 1944 | .740<br>1.940  | .360<br>.750   | 3.330<br>.320   | 3.940<br>0.000 | 9.280<br>0.000 | 1.000<br>0.000 |
| 1945 | 1.080<br>.280  | 3.240<br>.200  | 2.390<br>0.000  | 1.330<br>0.000 | 7.610<br>.240  | 4.880<br>0.000 |
| 1946 | 2.800<br>0.000 | 2.050<br>.430  | 7.690<br>0.000  | .850<br>0.000  | 2.770<br>0.000 | 3.280<br>0.000 |
| 1947 | .790<br>3.860  | .670<br>.620   | 2.360<br>.150   | .050<br>0.000  | 2.480<br>0.000 | 4.380<br>0.000 |
| 1948 | 1.740<br>.150  | .150<br>.510   | 6.010<br>0.000  | 1.180<br>.020  | 3.040<br>.030  | 5.960<br>0.000 |
| 1949 | .190<br>1.470  | 1.470<br>.260  | 1.690<br>0.000  | 6.090<br>0.000 | 2.970<br>.050  | 2.230<br>0.000 |
| 1950 | 2.720<br>1.370 | 6.330<br>.690  | 3.040<br>.050   | 2.950<br>0.000 | 1.940<br>0.000 | 1.040<br>.030  |
| 1951 | 1.440<br>.800  | 4.010<br>.220  | 8.360<br>0.000  | 9.800<br>0.000 | 1.610<br>0.000 | 6.820<br>.070  |
| 1952 | .070<br>1.890  | 3.000<br>.620  | 8.760<br>.100   | 2.680<br>0.000 | 0.000<br>.090  | 2.090<br>0.000 |
| 1953 | .300<br>.530   | 2.160<br>.190  | .490<br>.250    | 4.370<br>0.000 | 3.280<br>0.000 | 4.680<br>0.000 |
| 1954 | 0.000<br>2.840 | 1.980<br>1.360 | 2.980<br>0.000  | 5.530<br>0.000 | 1.990<br>0.000 | .310<br>0.000  |
| 1955 | .030<br>1.600  | 2.240<br>.480  | 14.940<br>0.000 | 6.590<br>0.000 | 2.250<br>0.000 | .510<br>.210   |
| 1956 | .790<br>1.720  | .020<br>1.930  | .530<br>.090    | 5.210<br>0.000 | 5.110<br>0.000 | 1.570<br>.070  |
| 1957 | 1.470<br>8.500 | .930<br>.560   | 3.990<br>.140   | 4.630<br>0.000 | 9.680<br>0.000 | 7.820<br>.780  |



| YEAR | OCT<br>APR     | NOV<br>MAY     | DEC<br>JUN     | JAN<br>JUL      | FEB<br>AUG      | MAR<br>SEP     |
|------|----------------|----------------|----------------|-----------------|-----------------|----------------|
| 1958 | .030<br>.110   | 1.010<br>.090  | .270<br>0.000  | 6.350<br>0.000  | 6.920<br>.090   | .230<br>2.100  |
| 1959 | 0.000<br>1.910 | 0.000<br>.330  | .710<br>0.000  | 5.260<br>0.000  | 6.520<br>0.000  | .740<br>.050   |
| 1960 | .050<br>.860   | 3.860<br>.520  | 1.810<br>.070  | 1.910<br>0.000  | .910<br>0.000   | 2.490<br>0.000 |
| 1961 | .170<br>.110   | 2.620<br>.340  | 1.300<br>0.000 | 2.000<br>0.000  | 11.290<br>0.000 | 2.960<br>0.000 |
| 1962 | 1.810<br>3.990 | .100<br>.230   | 2.290<br>.050  | 7.810<br>0.000  | 2.200<br>0.000  | 4.420<br>.090  |
| 1963 | 1.300<br>.390  | 3.640<br>1.350 | .460<br>.170   | 4.000<br>0.000  | .460<br>.300    | 3.080<br>0.000 |
| 1964 | 1.290<br>2.850 | 3.500<br>.140  | 4.830<br>.020  | 2.830<br>0.000  | 1.050<br>.250   | 2.520<br>0.000 |
| 1965 | .070<br>.270   | 6.630<br>0.000 | 4.790<br>.080  | 1.810<br>.290   | 1.340<br>0.000  | .680<br>.180   |
| 1966 | 0.000<br>6.410 | 3.740<br>.520  | 5.410<br>.190  | 6.090<br>0.000  | .430<br>0.000   | 6.160<br>.130  |
| 1967 | .270<br>.670   | 1.750<br>.430  | 1.870<br>.120  | 3.360<br>0.000  | 1.010<br>.020   | 3.290<br>0.000 |
| 1968 | .330<br>2.140  | 1.950<br>.120  | 4.060<br>.100  | 15.160<br>0.000 | 11.970<br>0.000 | 1.020<br>0.000 |
| 1969 | .340<br>.810   | .790<br>.070   | 3.110<br>.040  | 7.000<br>0.000  | 4.610<br>0.000  | 1.510<br>0.000 |
| 1970 | .110<br>1.210  | 5.860<br>.150  | 6.440<br>0.000 | 1.200<br>0.000  | .530<br>.040    | 1.490<br>.170  |
| 1971 | .240<br>.890   | 1.430<br>.080  | 5.560<br>.020  | 1.190<br>0.000  | .980<br>0.000   | .010<br>.030   |
| 1972 | 2.850<br>.100  | 5.790<br>.020  | 2.660<br>0.000 | 7.590<br>0.000  | 9.200<br>0.000  | 4.430<br>.070  |
| 1973 | 1.810<br>1.970 | 6.310<br>0.000 | 2.360<br>.110  | 4.440<br>.070   | 1.980<br>0.000  | 4.740<br>0.000 |



# Logged Anomalies

| YEAR | OCT<br>APR     | NOV<br>MAY     | DEC<br>JUN     | JAN<br>JUL     | FEB<br>AUG     | MAR<br>SEP      |
|------|----------------|----------------|----------------|----------------|----------------|-----------------|
| 1926 | -.093<br>-.090 | 1.275<br>-.073 | -.437<br>.003  | -.347<br>-.015 | .705<br>-.031  | -.372<br>.065   |
| 1927 | .144<br>-.144  | .170<br>-.271  | -.065<br>-.092 | -.952<br>-.015 | -.212<br>-.031 | .279<br>-.075   |
| 1928 | -.424<br>.011  | .455<br>-.320  | .076<br>.710   | -.768<br>-.015 | -.581<br>-.031 | .115<br>-.075   |
| 1929 | -.444<br>.029  | -.961<br>.639  | -.804<br>-.092 | .269<br>-.015  | .005<br>-.031  | .415<br>-.046   |
| 1930 | -.395<br>-.418 | -.053<br>.187  | -1.359<br>.244 | .077<br>-.015  | -.504<br>-.031 | -.591<br>-.075  |
| 1931 | -.424<br>-.686 | .050<br>-.013  | 1.013<br>-.092 | .089<br>-.015  | .294<br>-.031  | -.655<br>-.075  |
| 1932 | -.444<br>-.532 | -.796<br>.279  | .008<br>-.053  | .486<br>-.015  | -.792<br>-.031 | -.299<br>-.075  |
| 1933 | .308<br>-.695  | -.961<br>.295  | .817<br>.450   | -.160<br>-.015 | .395<br>-.031  | -1.243<br>-.017 |
| 1934 | -.304<br>1.116 | .397<br>-.320  | -.102<br>-.092 | .496<br>-.015  | -.873<br>.469  | .529<br>-.075   |
| 1935 | -.129<br>.424  | -.368<br>.072  | -.306<br>.200  | -.285<br>.343  | .951<br>-.031  | -.257<br>-.075  |
| 1936 | -.086<br>-.398 | -.961<br>-.320 | .233<br>.030   | .143<br>-.015  | .606<br>-.031  | .929<br>-.075   |
| 1937 | -.386<br>.472  | -.263<br>-.320 | .619<br>-.092  | -.092<br>-.015 | 1.196<br>-.031 | .964<br>-.075   |
| 1938 | .150<br>-.466  | -.248<br>-.121 | -.161<br>.003  | -.070<br>-.015 | -.136<br>-.031 | .163<br>.020    |
| 1939 | .269<br>-.347  | -.646<br>-.081 | -.190<br>-.092 | .739<br>-.015  | .854<br>-.031  | -.043<br>.195   |
| 1940 | .269<br>-.347  | -.646<br>-.081 | -.190<br>-.092 | .739<br>-.015  | .854<br>-.031  | -.043<br>.195   |
| 1941 | -.072<br>.988  | -.676<br>.156  | .938<br>.003   | .238<br>-.015  | .991<br>-.031  | 1.052<br>-.075  |



| YEAR | OCT<br>APR     | NOV<br>MAY     | DEC<br>JUN     | JAN<br>JUL      | FEB<br>AUG     | MAR<br>SEP     |
|------|----------------|----------------|----------------|-----------------|----------------|----------------|
| 1942 | .312<br>.960   | -.248<br>.378  | 1.033<br>-.092 | .321<br>-.015   | -.203<br>-.031 | .211<br>-.075  |
| 1943 | .110<br>-.090  | .124<br>-.320  | -.441<br>-.006 | .658<br>-.015   | -.215<br>-.031 | .313<br>-.075  |
| 1944 | .110<br>.261   | -.654<br>.239  | .057<br>.185   | .014<br>-.015   | .886<br>-.031  | -.550<br>-.075 |
| 1945 | .289<br>-.570  | .483<br>-.138  | -.187<br>-.092 | -.737<br>-.015  | .709<br>.184   | .529<br>-.075  |
| 1946 | .391<br>-.817  | .154<br>.037   | .754<br>-.092  | -.968<br>-.015  | -.117<br>-.031 | .211<br>-.075  |
| 1947 | .138<br>.764   | -.448<br>.162  | -.196<br>.048  | -1.534<br>-.015 | -.197<br>-.031 | .440<br>-.075  |
| 1948 | .564<br>-.677  | -.821<br>.092  | .539<br>-.092  | -.804<br>.005   | -.048<br>-.002 | .697<br>-.075  |
| 1949 | -.270<br>.087  | -.057<br>-.089 | -.419<br>-.092 | .376<br>-.015   | -.066<br>.017  | -.071<br>-.075 |
| 1950 | .870<br>.046   | 1.031<br>.204  | -.012<br>-.043 | -.209<br>-.015  | -.366<br>-.031 | -.530<br>-.046 |
| 1951 | .448<br>-.229  | .650<br>-.121  | .828<br>-.092  | .797<br>-.015   | -.485<br>-.031 | .814<br>-.007  |
| 1952 | -.376<br>.244  | .425<br>.162   | .870<br>.003   | -.280<br>-.015  | -1.444<br>.055 | -.115<br>-.075 |
| 1953 | -.181<br>-.392 | .189<br>-.146  | -1.009<br>.131 | .098<br>-.015   | .010<br>-.031  | .494<br>-.075  |
| 1954 | -.444<br>.528  | .131<br>.538   | -.027<br>-.092 | .293<br>-.015   | -.349<br>-.031 | -.973<br>-.075 |
| 1955 | -.414<br>.138  | .214<br>.072   | 1.361<br>-.092 | .444<br>-.015   | -.266<br>-.031 | -.831<br>.116  |
| 1956 | .138<br>.183   | -.941<br>.755  | -.983<br>-.006 | .243<br>-.015   | .366<br>-.031  | -.299<br>-.007 |
| 1957 | .460<br>1.434  | -.304<br>.124  | .199<br>.039   | .145<br>-.015   | .924<br>-.031  | .934<br>.502   |





| YEAR | OCT<br>APR     | NOV<br>MAY     | DEC<br>JUN      | JAN<br>JUL     | FEB<br>AUG      | MAR<br>SEP      |
|------|----------------|----------------|-----------------|----------------|-----------------|-----------------|
| 1958 | -.414<br>-.713 | -.263<br>-.234 | -1.169<br>-.092 | .412<br>-.015  | .625<br>.055    | -1.036<br>1.056 |
| 1959 | -.444<br>.251  | -.961<br>-.035 | -.872<br>-.092  | .251<br>-.015  | .573<br>-.031   | -.689<br>-.026  |
| 1960 | -.395<br>-.197 | .620<br>.098   | -.375<br>-.025  | -.515<br>-.015 | -.797<br>-.031  | .007<br>-.075   |
| 1961 | -.287<br>-.713 | .325<br>-.029  | -.575<br>-.092  | -.484<br>-.015 | 1.064<br>-.031  | .133<br>-.075   |
| 1962 | .589<br>.790   | -.866<br>-.113 | -.217<br>-.043  | .593<br>-.015  | -.281<br>-.031  | .447<br>.011    |
| 1963 | .389<br>-.488  | .574<br>.534   | -1.030<br>.065  | .026<br>-.015  | -1.066<br>.231  | .163<br>-.075   |
| 1964 | .385<br>.531   | .543<br>-.189  | .355<br>-.072   | -.240<br>-.015 | -.726<br>.192   | .015<br>-.075   |
| 1965 | -.376<br>-.578 | 1.071<br>-.320 | .348<br>-.015   | -.550<br>.240  | -.594<br>-.031  | -.724<br>.090   |
| 1966 | -.444<br>1.136 | .595<br>.098   | .450<br>.082    | .376<br>-.015  | -1.087<br>-.031 | .726<br>.047    |
| 1967 | -.205<br>-.304 | .050<br>.037   | -.354<br>.021   | -.110<br>-.015 | -.746<br>-.012  | .213<br>-.075   |
| 1968 | -.159<br>.327  | .121<br>-.207  | .213<br>.003    | 1.200<br>-.015 | 1.118<br>-.031  | -.540<br>-.075  |
| 1969 | -.151<br>-.224 | -.379<br>-.253 | .005<br>-.053   | .496<br>-.015  | .280<br>-.031   | -.323<br>-.075  |
| 1970 | -.339<br>-.024 | .965<br>-.180  | .599<br>-.092   | -.794<br>-.015 | -1.019<br>.008  | -.331<br>.082   |
| 1971 | -.229<br>-.181 | -.073<br>-.243 | .473<br>-.072   | -.799<br>-.015 | -.761<br>-.031  | -1.233<br>-.046 |
| 1972 | .904<br>-.722  | .954<br>-.300  | -.111<br>-.092  | .568<br>-.015  | .878<br>-.031   | .449<br>-.007   |
| 1973 | .589<br>.271   | 1.028<br>-.320 | -.196<br>.012   | .111<br>.053   | -.352<br>-.031  | .504<br>-.075   |



# Reserved Data

| YEAR | OCT<br>APR     | NOV<br>MAY    | DEC<br>JUN     | JAN<br>JUL     | FEB<br>AUG      | MAR<br>SEP     |
|------|----------------|---------------|----------------|----------------|-----------------|----------------|
| 1974 | 1.840<br>1.660 | .760<br>.020  | 3.570<br>.040  | 2.040<br>.070  | 5.490<br>.120   | 0.200<br>0.000 |
| 1975 | 1.310<br>1.460 | .730<br>.050  | .330<br>.180   | .080<br>.001   | 2.350<br>1.050  | 1.450<br>.750  |
| 1976 | 1.450<br>.140  | .600<br>.970  | 1.850<br>.080  | 2.450<br>0.000 | .460<br>0.000   | 1.800<br>.480  |
| 1977 | .020<br>3.600  | .400<br>.030  | 5.680<br>0.000 | 9.990<br>0.000 | 8.400<br>0.000  | 7.070<br>.360  |
| 1978 | 0.000<br>.380  | 2.720<br>.260 | 1.200<br>0.000 | 5.810<br>.120  | 4.660<br>0.000  | 3.920<br>0.000 |
| 1979 | 1.010<br>2.550 | 2.140<br>.420 | 3.870<br>0.000 | 5.220<br>.950  | 10.890<br>0.000 | 2.890<br>0.000 |



## LIST OF REFERENCES

1. Box, E.P., and Jenkins, G.M., Time Series Analysis: Forecasting and Control, 2d. ed., Holden-Day, 1970.
2. Gaver, D.P., Logistic Analysis, class presentation notes for OA4910, Naval Postgraduate School, Fall 1980.
3. Fleiss, J.L., Statistical Methods for Rates and Proportions, Wiley, 1973.
4. Fisher, R.A., The Design of Experiments, 7th. ed., Oliver and Boyd, 1960.
5. Dixon, W.J., and Massey, F.J. Jr., Introduction to Statistical Analysis, 3d ed., McGraw-Hill, 1969.
6. Brownlee, K.A., Statistical Theory and Methodology in Science and Engineering, 2d. ed., Wiley, 1965.
7. Mood, A.M., Grayhill, F.A., and Boes, D.C., Introduction to the Theory of Statistics, 3d ed., McGraw-Hill, 1963.
8. Nelson, C.R., Applied Time Series Analysis for Managerial Forecasting, Holden-Day, 1973.
9. Pierce, D.A., "A Survey of Recent Developments in Seasonal Adjustment", The American Statistician, Vol. 34, No. 3, August 1980.
10. Kilmartin, R.F., and Peterson, J.R., "Rainfall-Runoff Regression with Logarithmic Transforms and Zeros in the Data", Water Resources Research, Vol. 8, No. 4., August 1972.
11. Hipel, K.W., BcBean, E.A., and McLeod, A.I., "Hydrologic Generating Model Selection", Journal of the Water Resources Planning and Mangement Division, W.R. 2, September 1979.
12. Naval Postgraduate School Report NPS55-78-034, An Interactive Software Package for Time Series Analysis, by F. Russell Richards and Stephen R. Woodall, November 1978.
13. Mosteller, F., and Tukey, J.W., Data Analysis and Regression; A Second Course in Statistics, Addison-Wesley, 1977.



14. Kendall, M., Time-Series, 2d. ed., Hafner Press, 1976.
15. McNeil, D.R., Interactive Data Analysis, A Practical Primer, Wiley, 1977.





# INITIAL DISTRIBUTION LIST

|  | No. Copies |
|--|------------|
| 1. Defense Technical Information Center<br>Cameron Station<br>Alexandria, Virginia 22314                   | 2          |
| 2. Library, Code 0142<br>Naval Postgraduate School<br>Monterey, California 93940                           | 2          |
| 3. Professor D.P. Gaver, Code 55Gv<br>Naval Postgraduate School<br>Monterey, California 93940              | 1          |
| 4. Professor P.A. Jacobs, Code 55Jc<br>Naval Postgraduate School<br>Monterey, California 93940             | 1          |
| 5. Professor R.J. Renard, Code 63Rd<br>Naval Postgraduate School<br>Monterey, California 93940             | 2          |
| 6. Monterey Water Management District<br>Attn: Mr. Bruce Buel<br>P.O. Box 85<br>Monterey, California 93940 | 2          |
| 7. Professor F.D. Faulkner, Code 53Fa<br>Naval Postgraduate School<br>Monterey, California 93940           | 1          |
| 8. Professor A.L. Schoenstadt, Code 53Zh<br>Naval Postgraduate School<br>Monterey, California 93940        | 1          |
| 9. Professor K.T. Marshall, Code 55Mt<br>Naval Postgraduate School<br>Monterey, California 93940           | 1          |
| 10. Captain David F. Davis, USA<br>17 Glenmore Rd.<br>Pueblo, Colorado 81001                               | 1          |











192597

Thesis  
D169555  
c.1

Davis

A statistical analysis of monthly rainfall for Monterey Peninsula and the Carmel Valley in Central California.

27 JUN 84

29243

192597

Thesis  
D169555  
c.1

Davis

A statistical analysis of monthly rainfall for Monterey Peninsula and the Carmel Valley in Central California.

A statistical analysis of monthly rainfall



3 2768 001 02329 4

DUDLEY KNOX LIBRARY